Economics of professional football

Besters, Lucas

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Besters, L. (2018). *Economics of professional football*. [Doctoral Thesis, Tilburg University]. CentER, Center for Economic Research.

Economics of Professional Football

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 26 januari 2018 om 10.00 uur door

Lucas Mattheus Besters

geboren op 22 februari 1989 te Hulst

Promotiecommissie:

Promotor:	Prof. dr. ir. Jan van Ours
Copromotor:	Dr. Martin van Tuijl
Overige leden:	Prof. dr. Harry van Dalen
	Dr. Patricio Dalton
	Prof. dr. Ruud Koning

Acknowledgements

On the 26th of January 2012, I obtained my Master degree in Economics at Tilburg University, with Martin van Tuijl as supervisor and Jan van Ours and co-reader. Exactly six years later, on the 26th of January 2018, I defend this dissertation with Jan as promotor and Martin as co-promotor. Within these six years, I first worked as an entrepreneur for three and a half years. After that, I was given the opportunity to earn a PhD by writing four papers on the economics of professional football in only a short period of time. Thus, I have been working hard on this dissertation for the last two and a half years. Quite some people asked me whether I would be able to finish in time. To be honest, there were moments that I had doubts myself. However, I usually answered that my supervisors are confident about it, and, thus, so am I. Often, I added that it would not be possible without the extraordinary supervision of my promotors Jan and Martin. This is true without any doubt and I thank both of you for that. Our cooperation has always been very pleasant and I really appreciate all your advice. It is nice to have the feeling that you, as supervisors, are concerned about my interest during the writhing of this dissertation. Furthermore, I really enjoyed our conversations about our common hobby, i.e. football.

I remember that this already happened the first time that Jan and I met in his office in Tilburg during the summer of 2010. You were looking for a research assistant and Martin mentioned me as a candidate. We briefly discussed the tasks that I should perform. Then, we continued to talk about my experience as a football player and you being a fan of Feyenoord. Once, you invited me to join you to attend a home match in De Kuip. Although I am not a fan of Feyenoord myself, I really enjoyed that evening, and I must agree with you that the atmosphere in the stadium is special. In the past two years, we also attended two conferences on sports economics together. In the evenings, we had nice discussions about many things while tasting some beers. I think that it is a good sign that these moments are on top of my mind when I think of our cooperation, for which I want to thank you. However, I am most thankful for all the advice that you gave me and all your experience, expertise and knowledge that you shared with me. I learned a lot from you during the past couple of years. By working under your supervision, I became a better researcher and applied economist.

The same applies for my cooperation with Martin. You introduced me to the of field sports economics as the topic for my bachelor thesis in 2009. After that, you also supervised my master thesis and advised me during the writing of this dissertation. In all these years, we had numerous, unscheduled, long and enjoyable discussions about football. Especially the way in which you are able to translate all your knowledge of the sport into funny anecdotes, is a pleasure to listen to. Most of our conversations started with a question about data related to one of our present studies. While talking, we spent most of the time on other topics, such as the current situation in Dutch professional football, and potential research projects. At the end of our meetings, you usually came back to the initial question and then you advised me what to do. That worked perfectly for me. Furthermore, you provided me with advice and support on issues that influenced my writing process, so that I would be able to finish in time. I appreciate your critical view on my work. That definitely improved the quality of my dissertation and my scientific skills. I thank you for that.

I also want to thank Harry van Dalen, Patricio Dalton and Ruud Koning for being a member of the promotion committee and all your useful comments. I hope you enjoyed our discussion as much as I did. Your input especially helped me to rethink and improve the interpretation of my findings, which enhanced the quality of the chapters. Furthermore, I thank the participants of the two conferences organized by the European Sport Economics Association as well as the participants of the workshop on Economics and Management of Professional Football at the Erasmus University for their feedback on my presentations. Specifically, I want to mention Thomas Peeters and Gerard Sierksma. I also thank ORTEC Sports and, in particular, Bertus Talsma for the provision of data for, and some useful comments on, the fifth chapter of this doctoral thesis.

Naturally, I would not have been able to write this dissertation without the help and support of my close family and friends. First, I thank my parents, Cees and Marijke, for the love that they gave me throughout my life. You taught me that, if you want to achieve something, you should put effort in that. At the same time, you kept stressing that one cannot do more than the best you can. You challenged me to be curious and to learn new things, without putting any pressure on me. If you saw that I was too busy with my work, you remembered me of the importance to take some rest. I hope that you are proud of my achievements and, more importantly, you are proud of me as a person and as your son. I thank you for all your support and I am grateful to have you as my parents.

Special thanks go to my brother Casper, with whom I have been sharing an apartment for the past three years. Besides brother and roommate, you are a very good friend, with whom I can discuss any of my concerns. You were helpful and understanding in periods that I was very busy, while always in the mood to talk about my research. Based on your layman’s knowledge as a football fan, you provided me with insights I did not think of. As a joke, you, then, asked to become a co-author, and I smiled and answered that I would take that into consideration. In the end, I can only give you these well-deserved acknowledgments. I would say that we had quite a lot of fun together. Every now and then, your girlfriend Marjolein joint us during dinners and drinks. I thank you for these wonderful

moments. I also thank my other brother Victor and his girlfriend Lucienne for your support and advice whenever I needed. We share quite an experience and I am very happy to have you in my life.
Furthermore, I thank Matthijs and Stephanie for being close friends and all the moments that we spent together at the university. We had plenty of lunches, coffee-breaks and Friday-afternoon drinks, where we agreed not to discuss our research, since we already had to think of that for the rest of the day. Although that meant that I was not allowed to talk about football, I really enjoyed these moments and thank you for that. I also thank Constant for the fun we had during our weekly dinners at the university, my office-mates Jakub, Michela and Yuxin and my other fellow PhD students Shuai and Khulan. In addition, I thank Jacco, Judith and Ruud, with whom I spend a lot of time in Tilburg during the past couple of years.
Finally, I want to thank my friends Jelle, Anne and Vera, and the members of De Ronde Tafel: my brother Casper, Harm, Jason, Kim, Matthijs, Remco, Tias and Tom for all the fun moments we had.

Lucas M. Besters
Tilburg, December 2017

Contents

List of Figures

List of Tables

Chapter 1

Introduction

People have been interested in sports for a very long time. Predominantly, because it is an enjoyable leisure activity. Both, in an active way as a participant and more passive as a spectator. Scholars from the domain of economics are no exception in that respect. However, economists have become increasingly interested in sports from a professional point of view during the past few decades. First, because it has proven to be an interesting industry to study in itself. Second, because sports provide unique possibilities to study phenomena of interest to various domains, such as labour markets (Kahn, 2000).

This dissertation contains four chapters, all with a different topic that is of interest from a sports economic perspective. More specifically, from the economic perspective of professional football. Football is the most popular sport within Europe and the data that is used in the analyses stems from English and Dutch professional football. The topics also relate to elements outside of the sports domain. For example, the effectiveness of in-season coach changes shows resemblance to managerial changes within organisations (Chapter 2). Stadium attendance demand relates to the entertainment industry and describes consumer preferences in uncertain situations (Chapter 3). Furthermore, the selection system for talent in youth professional football is comparable to other selections system, such as in school grades. Both have to deal with relative age differences between peers (Chapter 4). Finally, the effects of team heterogeneity on performance relate to organisational structures and, specifically, the formation of teams (Chapter 5).

In Chapter 2 the analysis deals with performance effects of in-season manager changes within the English Premier League football during the seasons 2000/01 – 2014/15. It follows that some managerial changes are successful, while others are counterproductive. On average, performance does not improve following a managerial replacement. The development of performance around the time of the change in manager is subject to regression to the mean. Case studies illustrate that the successfulness of managerial turnover depends on specific highly unpredictable circumstances.

Chapter 3 investigates the determinants of stadium attendance in the highest level of Dutch professional football for the seasons 2000/01 – 2015/16. Since attendance rates are high in the Dutch *Eredivise*, with about 40 percent of the matches that can be considered as sold out, a Tobit model is used for the within-season variation in attendance rates. On average, attendants, i.e. the consumers in this setting, have reference-dependent preferences with loss aversion that dominate their preference for uncertain outcomes. This contradicts with the well-known uncertainty of outcome hypothesis (UOH). However, in general, team characteristics seem more important for the determination of stadium attendance than behavioural economic explanations regarding the outcome of the match. For seasonal uncertainty, many results are in line with the UOH. Moreover, the introduction of play-offs in the season 2005/06 has had a positive effect on stadium attendance during regular league matches. Although these results are statistically significant, the economic impact in terms of additional attendance is small. As to the between-season variation, we find a high positive correlation between stadium attendance and stadium capacity, suggesting excess demand for tickets.

In Chapter 4, I look at the selection of talent in relation to the presence of a relative age effect (RAE). Many selection systems suffer from a bias with respect to the selection of players who are born just posterior to the cut-off date. The skewed birth-date distribution that results, with an overrepresentation of early-born and age-advantaged players, is known as the RAE. Under the assumption that talent is uniformly distributed across birth dates, this suggests that talent is lost. With data from PSV Eindhoven (PSV), it follows that an RAE is persistent within their youth academy. Furthermore, I show that this results from external selection, i.e. the recruitment of players from outside of the academy. Internal selection, i.e. the annual decision whether players may stay or have to leave, reduces the severity of the RAE. Finally, most of the players who eventually become a professional football player, are early-born. However, at the age of 19, late-born players have a higher probability to become a professional. This suggests that only the highly talented late-born players are selected. The underlying assumptions as well as the generalizability of the results are extensively discussed.

The final Chapter 5 provides a study on the relationship between team heterogeneity and performance with data from the highest tier of Dutch professional football in the season 2014/15. Performance is measured by a unique Success Ratio of individual player actions. This measure is used in both, an individual player analysis as well as in a team level analysis. The results reveal that heterogeneity concerning nationality influences performance positively, whereas heterogeneity with respect to experience affects performance negatively. Although these results are statistically significant, their economic impact in terms of performance differences is small. Furthermore, heterogeneity related to ability and heterogeneity related to height are insignificant. For team performance, it follows that a continuous measure of performance, based on a team average of individual Success Ratio's, might better reveal certain relationships than a discrete measure that is based on match outcomes.

Chapter 2

Effectiveness of in-season manager changes in English Premier League football

(Published as: Besters, L.M., Van Ours, J. C., and Van Tuijl, M. A. (2016) Effectiveness of in-season manager changes in English Premier League football, *De Economist*, 164, 335-356.)

2.1 Introduction

Football is very popular worldwide. In Europe and Latin-America, football has entertained crowds for more than one century. In other continents, interest has increased in the past decades. Top players now move to the football leagues of Australia, Japan and the United States, and, more recently, also to the league of the People's Republic of China. Both clubs and national associations employ top-class managers from all around the globe to coach their squads. Furthermore, top clubs have an enormous global fan base.

The great interest in football is not restricted to fans seeking entertainment. Professional sports, in general, and professional football, in particular, have proven to be a fruitful soil for scientific research. Kahn (2000) and Szymanski (2003), for example, argue that professional sports offer interesting data to analyse labour market phenomena. In this respect, the high frequency of data obtained from controlled events is of particular interest. Results of football matches, for example, provide a straightforward and objective measure of performance (Ter Weel, 2011). An element in football that has clear analogies both with business and economics is the ongoing debate about the effects of management on the performance of firms. Kuper and Szymanski (2010) question the influence of managers on the performance of professional football teams. In contrast, Anderson and Sally (2013) argue that this influence is non-negligible, as leadership appears to matter for history, in general, and for business, in particular. Pieper, Nüesch and Franck (2014) argue, that football managers closely resemble managers in other branches of the economy with respect to personal characteristics, such as age and the capabilities to cope with stress, media attention and a large group of stakeholders.

Assuming that 'management matters', or, at least, that decision-makers, such as the supervisory board, suppose that 'management matters', one of the key decisions is the hiring and firing of managers (Ward et al., 2011). This holds in particular, if the decision to sack a manager is implemented prior to the expiration of her/his contract. Summary dismissals tend to be rather costly, so that one aims for improved performances in return. Both Pieper, Nüesch and Franck (2014) and Van Ours and Van Tuijl (2016) show that the decision to replace a manager is related to the difference between actual performance and expectations. Both studies use bookmaker odds to derive these expectations. The authors find an increased probability of replacement, if actual performances fall short of expectations. Thus, a sequence of (rather) bad results triggers clubs to replace the manager, hoping for better performances afterwards (Bruinshoofd and Ter Weel, 2003).

Much research has already been done on the effects of manager turnover in business. These studies mainly use stock prices, or data derived from financial statements that are only published with a lag, viz. on a quarterly or annual basis. These outcomes point at a statistically significant but small positive effect (Ter Weel, 2011). Studies on the effectiveness of managerial changes in professional football have been done for a variety of European countries, for example, Belgium, England, Germany, Italy, the Netherlands and Spain (see for a recent overview Van Ours and Van Tuijl (2016)).

Two Belgian studies, Balduck *et al.* (2010a) and Balduck *et al.* (2010b), find no performance effects of a coach replacement. Studying English football, Poulsen (2000) finds no effects of a managerial change while Dobson and Goddard (2011) find a negative effect, just after the replacement of a manager. Analyzing data from German football, Salomo and Teichmann (2000) find negative effects of a trainer-coach dismissal, while Hentschel *et al.* (2012) conclude that a coach change may have a positive effect on homogeneous teams but no effect for heterogeneous teams. De Paola and Scoppa (2011) find similar conclusions for Italian football, just like Tena and Forrest (2007) for Spanish football. Koning (2003), Bruinshoofd and Ter Weel (2003), Ter Weel (2011), Van Ours and Van Tuijl (2016) study the effects of the replacement of head-coaches in the highest professional football league of the Netherlands. They all find that this does not lead to better performance of the teams involved.

We study the effects of managerial changes using data of the English Premier League. We apply the method initially used by Van Ours and Van Tuijl (2016). Studying Dutch professional football, they account for potential selectivity of managerial changes by, first, correcting for the strength of the opponent and, second, by defining a counterfactual case with a similar development of performances prior to the hypothetical change, but without the managerial change actually taking place. The authors use the so-called cumulative surprise as an indicator of the difference between performance and expectations. The cumulative surprise is the sum of the differences between the actual number of points and the expected number of points, as based on bookmaker odds. Then, they use this cumulative surprise to match actual coach changes to counterfactual observations. In line with most previous studies, Van Ours and Van Tuijl (2016) conclude that the development of performances around the time of the change in trainer-coach is subject to regression to the mean.

Our main finding is that, on average, an in-season replacement of the manager has no effect on in-season performances. In addition to the replication of the method of Van Ours and Van Tuijl (2016) for the English Premier League, we also investigate whether there is heterogeneity in the effects and find that some changes have positive effects, while other changes are counterproductive, i.e. the effects of a managerial replacement on team performance are negative. To find out whether there is a pattern in this heterogeneity of the effects of a managerial change, we also study subsamples. These subsamples are based on the origin of the manager (British versus non-British), his age, whether or not the manager ever played for a national team, whether the team was recently promoted to the Premier League and whether the team finished top-10 or bottom-10 in the preceding season. Our main finding, i.e. managerial replacements are ineffective, stands up to the scrutiny of these subsamples. To explore potential differences between successful and unsuccessful managerial changes we present three case-studies, from which we conclude that the efficacy of managerial turnover depends on specific highly unpredictable circumstances.

Our paper is organized as follows. In section 2.2, we present our data and our research method. Subsequently, we discuss our results in section 2.3. Next, we present three case-studies in section 2.4. Finally, section 2.5 concludes.

2.2 Data and set-up of the analysis

We use data from English Premier League (EPL) football for 15 seasons, from 2000/01 to 2014/15. Every season contains 20 clubs that compete according to a double round-robin format, resulting in 380 matches per season (5,700 matches in total). For every match, the date, the home team, the away team and the final score are recorded. Furthermore, the dataset contains match-specific bookmaker data concerning the final result, as well as the managers in charge of the two teams per match.[1] Thus, information on in-season changes is included.[2] In case of a managerial change, we distinguish between forced 'sackings' and voluntary 'resignations'.[3] Finally, the dataset contains information on the final ranks of all clubs within the EPL in the preceding season.

In our analysis, we consider the first managerial change of a club within a particular season. Thus we ignore, for example, a caretaker who is replaced after some matches by a newly hired manager.

[1] The bookmaker data stem from William Hill (98 per cent) and from Ladbrokes (two per cent), in case WH data were lacking.

[2] The data on managers has mainly been collected from *www.soccerbase.com*. In case of missing or ambiguous information, we have examined newspaper archives and other internet sources.

[3] We only consider the first managerial change of a particular club within a particular season. We have collected information on sackings and resignations from newspaper archives and BBC Sport.

Consequently, the sample period contains 84 in-season managerial changes. We follow Van Ours and Van Tuijl (2016) in their method of analyses. They discuss coach changes in two steps. First, they show that the probability of a coach change depends on the in-season performance of the team. The in-season performance is measured by the number of points in the last four matches as well as the cumulative surprise, i.e. the cumulative difference between the expected number of points and the actual number of points obtained. Expectations are based on bookmaker odds. The second step, which is replicated in the present study, is to test the performance effects of the coach changes. The development of performances before and after the change is compared with the development of performances in case the change would not have happened. Since the latter cannot be observed, there is a need to construct a control group. In order to be a valid counterfactual, an observation needs to fulfil the following five requirements:

1. The observation concerns the same club, but stems from a different season that does not contain an in-season change in manager. This excludes two types of changes. First, we ignore changes that occurred at clubs that only played in the EPL during just one season in the sample period. Second, we do not take changes into account at clubs that changed their manager in all of their EPL-seasons in the sample period.
2. The observation should exhibit a cumulative surprise that does not differ more than 0.5 from the cumulative surprise at the time of the actual managerial change. This leads to the exclusion of cases that exhibit a rather large (positive or negative) cumulative surprise at the time of the change, compared to all other observations. Applying such a maximum value potentially results in the exclusion of both rather successful cases and rather unsuccessful cases.[4]
3. Consistency with the actual managerial changes requires that we exclude matching with an observation prior to the fifth match and posterior to match 34.
4. For observations that fulfil the first three requirements, we look for the smallest difference between the rank number of the last match of the replaced manager and the rank number of the match attached to the potential counterfactual. By doing so, we assure that matching is also based on the time during the season at which a change takes place. The closer the rank numbers of the matches, the higher the likelihood that the pattern towards the change is similar as compared to the counterfactual. Furthermore, it makes sure that the performances of the

[4] Obviously, this value of 0.5 is fairly arbitrary. Yet, an extensive sensitivity analysis has made clear that different values only lead to a small change in the number of cases to be considered, without altering the main conclusions.

treatment group and the control group have a more or less similar period (i.e. in terms of the number of matches or observations) to develop, after the treatment has taken place.

5. In case multiple observations meet all previous requirements, we take as the counterfactual observation the one with the smallest difference in cumulative surprise as compared to the actual observation.

The idea is to find a situation (i.e. season) without coach change that is comparable to the one in which the coach has been changed. Since club-specific elements might matter for the hiring and firing of coaches, such as the sentiment of fans, we want the counterfactual to come from the same club. Independence between observations requires that the counterfactual does not contain a coach change itself. Resemblance of the counterfactual observation with the actual coach change is primarily obtained through resemblance of the cumulative surprise. Thus, the in-season performance of the control group should correspond with the in-season performance at the time of a coach change. Furthermore, since the cumulative surprise includes expectations that capture and control for season-specific quality etc., it is safe to compare performances between seasons. Then, since we compare performances before and after the (hypothetical) change, we want the number of matches before (after) the actual change to be close to the number of matches before (after) the counterfactual change. This allows for rather equal opportunities in the development of performances and, thus, a comparable situation.

In our sample, ten of the managerial changes occurred, either prior to the fifth match or posterior to the 34th match. These 10 changes will be left out of our analysis. Thus, 74 managerial changes remain, of which 13 do not meet the criteria for matching with a valid counterfactual case to be used in the difference-in-differences approach. Our final sample thus consists of 61 managerial changes. Table 2.1 presents descriptive statistics of this sample. It shows, per season and in total, the number of changes to be considered in our analysis for the complete sample, dismissals only and the subsamples that are based on managerial characteristics. We define 'British' managers as managers from the United Kingdom and from the Republic of Ireland, thus making a sharp distinction between these two countries and the rest of the world. For age, we distinguish between managers aged over 50 and managers aged under 50 at the time of replacement. The age of 50 is the overall average and splits the sample in two more or less equal subsamples. Table 2.1 further shows that managerial replacements, on average, take place around the middle of the season (column 'W') with the cumulative surprise then being negative (column 'CS'). Column 'FS' shows the average cumulative surprise at the end of the season, indicating that, for some seasons we find improvements, while for others the cumulative surprise decreases. The last three columns show the average values for our counterfactual observations. By definition, the values for the rank number of the match and for cumulative surprise are rather similar for the treatment and control group. However, the improvement in cumulative surprise is larger for the control group than for the treatment group, given the values for MFS and FS. Table A1 in the appendix presents a detailed overview including all single managerial changes.

Table 2.1: Descriptives

Season	Changes	D	B	A	Age	C	W	CS	FS	MW	MCS	MFS
00/01	4	3	3	1	45.0	4	19.7	-2.1	-1.1	24.2	-1.8	-2.3
01/02	5	3	5	2	50.6	2	15.0	-2.2	-4.0	15.0	-2.3	-1.7
02/03	2	2	1	0	47.1	2	21.0	-2.6	-9.7	21.5	-2.8	-1.0
03/04	3	2	3	0	46.8	3	14.3	-3.3	-5.0	15.3	-3.1	4.4
04/05	6	2	5	5	55.4	3	11.8	-2.3	-3.4	12.5	-2.3	0.0
05/06	2	2	1	1	50.9	1	18.0	-4.5	1.3	24.5	-4.3	-2.2
06/07	2	2	2	0	41.1	1	25.0	-5.1	-4.6	25.0	-5.2	-1.0
07/08	7	4	6	2	47.9	2	13.1	-4.4	-5.2	13.9	-4.3	0.4
08/09	4	2	3	2	50.1	2	16.2	-2.3	-3.5	13.2	-2.4	-4.7
09/10	4	4	4	3	51.0	1	20.2	-3.2	-3.7	22.5	-3.4	-1.5
10/11	5	4	4	3	52.4	2	16.6	-2.3	-0.8	18.6	-2.5	-5.1
11/12	2	2	1	1	42.7	0	20.0	-6.0	-4.8	21.5	-5.9	-5.0
12/13	4	4	3	2	50.8	2	23.5	-2.0	-1.9	18.7	-2.3	-0.2
13/14	7	6	4	5	49.5	3	17.9	-3.4	-1.5	15.0	-3.6	-2.4
14/15	4	3	3	2	50.7	2	23.0	-2.0	-2.1	23.0	-2.0	-1.6
Total	61	45	48	29	49.6	30	17.5	-3.0	-3.1	17.7	-3.0	-1.5

Note: ‘Changes’ indicate the number of changes (all changes included in the analyses) while ‘D’ is the number of dismissals, ‘B’ the number of British managers, ‘A’ the number of managers aged above 50, ‘Age’ is the average age at the time of replacement, ‘C’ the number of capped managers, ‘W’ the average number of the last match of the manager, ‘CS’ the cumulative surprise at the time of replacement, ‘FS’ is the average final surprise (at the end of the season) for teams that replaced their manager. The ‘M’ in the last three columns indicate that these values belong to the matched observations.

We estimate the parameters of the following linear model using OLS:

$$y_{ijk} = \eta_{ik} + r'_{ijk}\beta + \delta d_{ijk} + \lambda c_{ijk} + \varepsilon_{ijk}, \tag{1}$$

where y_{ijk} is the performance indicator, i denotes the club, j indicates the match and k refers to the season. We use the number of points as performance indicator.[5] Note that we investigate in-season replacements and performances. Therefore, we include club-season fixed-effects η_{ik}, which account for unobserved elements such as the quality of a team in a particular season. Home advantage is highly relevant for the performance (see for example Van Ours and Van Tuijl (2016)). Consequently, a dummy is included that has value one for matches played at home. Evidently, the quality of the

[5] Alternatively, we used victory (whether a team has won the match) and goal difference as performance indicators. Then, our main findings are identical.

opponent is also important. This strength is proxied by the final rank in the previous season.[6] The latter two variables are both included in the vector r'_{ijk}, while β represents the vector of parameter estimates and ε_{ijk} is the error term. The focus of our analysis is on two variables. First, d_{ijk} is a dummy for the treatment group, with value one if a manager has been replaced and δ measuring the effect of the managerial change on the performance. Second, c_{ijk} is a dummy for the control group, with value one if the 'hypothetical' change has taken place and with λ measuring the counterfactual effect on the performance. An F-test for the equality of δ and λ reveals whether the managerial change exerts influence on the in-season performance. First, we estimate the parameters of equation (1) using our complete sample. Then, we estimate the relevant parameters for dismissals only.

Figure 2.1 shows kernel densities for the cumulative surprises at the end of the season for the subsets of dismissals, resignations, as well as for the majority of the cases, in which no managerial change has taken place. The distribution of the cumulative surprise for the dismissals is somewhat different from the distribution of the cumulative surprise for quits. Nevertheless, they look fairly similar. However, there is a clear difference between the seasons with a managerial change compared to the seasons without a managerial change. At the end of the latter seasons there is a more positive cumulative surprise. In other words, seasons with managerial changes are seasons with worse performance than seasons without a managerial change.

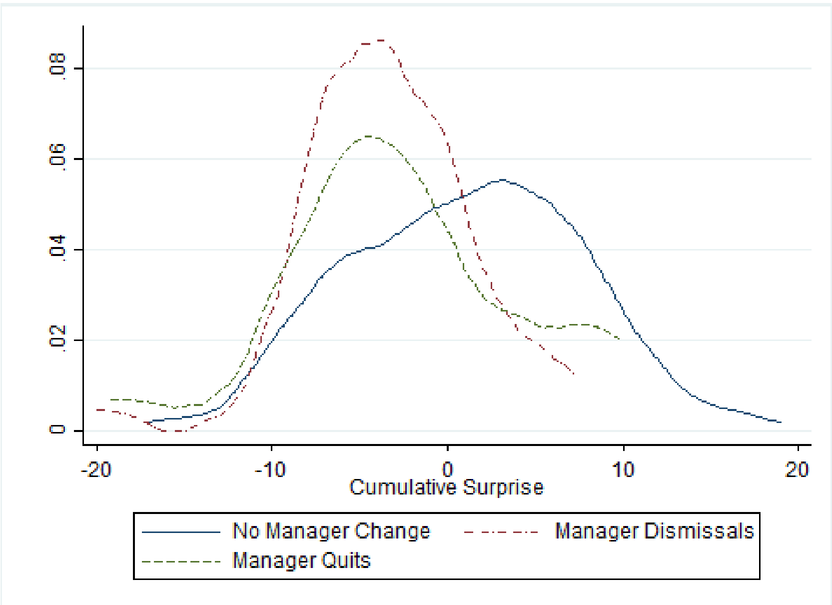


Figure 2.1: Kernel Density Cumulative Surprise for types of managerial changes; final match of the season

[6] Promoted teams are all assigned rank 20.

2.3 Parameter estimates

In analysing the effectiveness of in-season manager replacements, we first use all 61 changes for which we have found a valid counterfactual. Then, we focus on the subset of dismissals. The parameter estimates for all managerial changes are presented in the first columns of Table 2.2. ‘Rank Opponent’ is a measure of the strength of the opponent, while ‘Home’ represents home advantage. The variable ‘Manager change’ measures the difference in performance before and after a managerial change. Without taking a control group into account, we can interpret the coefficients of this variable as treatment effects. A significant positive value indicates that performances improved, suggesting that changes were effective. However, interpreting this result as causal would be wrong, since one does not take into account the situation in which the manager would not have been replaced. Therefore, we include a dummy variable for the control group reflecting managerial replacement that did not take place. Significant and positive values for the related parameter indicate that performances went up after the ‘counterfactual’ change, i.e. the matched observation. The *F*-test for equality between the two managerial-change parameter shows whether there is indeed a causal effect, i.e. if the two parameters are not significantly different from each other there is no treatment effect. Table 2.2 also shows the number of observations in the treatment and control groups, both separately and combined. Differences in the number of observations between the treatment and control groups arise because some club-season combinations are a control group for multiple treatment groups.

Table 2.2: Parameter estimates determinants team performance

	All changes	Dismissals
Rank Opponent	0.05***	0.05***
	(0.00)	(0.00)
Home	0.56***	0.55***
	(0.04)	(0.04)
Manager change	0.21***	0.28***
	(0.05)	(0.05)
Counterfactual	0.21***	0.26***
manager change	(0.06)	(0.06)
F-test for equality	0.00	0.03
Observations	4,028	3,002
n-Seasons	106	79
n-Treatment-Group	61	45
n-Control-Group	45	34

Note: Team performance is measured by the number of points per match. Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All estimates include club-season fixed effects. ‘Rank opponent’ refers to the rank of the opponent in the preceding season. ‘Home’ indicates whether a match was played at home.

Interpreting Table 2.2, while focusing on the results for all changes, we observe that both the strength of the opponent as well as the home advantage are highly significant. They both have the expected sign. The weaker the opponent, the better the outcome, while home matches also result in better results. Furthermore, we find a highly significant and positive coefficient for a managerial change. Our naïve conclusion would be that a change in manager is successful on average. However, we find similar results and comparable values for the counterfactual managerial change. The *F*-test indeed shows that there is no significant difference between the treatment and control group. The results thus show that the improvement in performance after the change in manager (i.e. the treatment group) would also have occurred if the manager would have kept his position (i.e. the control group). On average, we do not find a causal relation between performances and the managerial changes. This finding is in line with the results of previous studies and in particular comparable to the results found by Van Ours and Van Tuijl (2016).

The findings for dismissals are fairly similar. These results are presented in the second column of Table 2.2. Leaving out the 16 resignations, thus analysing 45 dismissals, results in comparable values, significance and conclusions. In general, thus, we may conclude that there is no point in firing a manager after a sequence of bad results, since performances would have improved irrespective of the manager in charge.[7] Again, these results are in line with previous studies.

Table 2.3 shows the results for multiple subsamples which are based on the characteristics of the replaced manager.[8] In the first and second column, we distinguish between British ($n=48$) and non-British ($n=13$) coaches.[9] Column three and four contain the results for subsamples of coaches aged over 50 ($n=29$) and aged under 50 ($n=32$), at the time they were replaced. Finally, the last two columns, five and six, report the results for those coaches who were capped as an active player ($n=30$) and those who did not play for their country ($n=31$). Without going into detail, the general result is that we find significant improvements in performance after a managerial change, which is also the case for the counterfactual managerial change. However, we do not find any significant differences

[7] It would be more accurate to formulate 'after a sequence of results below expectations', which emphasizes that clubs (probably) take into account the heterogeneity of opponents and the order of play in their decision to fire a manager. From Table A1 in the appendix it becomes clear that the cumulative surprise at the moment of the managerial replacement is negative for most cases.

[8] Note that for each group of two subsamples (i.e. British, Age and Capped) the total number of treatment groups is 61 and equal to the number for all changes in Table 2.2. However, the total number control groups might be different and in particular higher than the number of 45 in Table 2.2, since a club-season that is a counterfactual for multiple treatment groups is counted only once in Table 2.2, but twice if it belongs to both subsamples per group in Table 2.3.

[9] As mentioned above, we define 'British' managers as managers from either the United Kingdom or from the Republic of Ireland.

between treatment group and control group. This leads us to conclude that on average, for none of the subsamples, performances improve after a managerial change.

Table 2.3: Parameter estimates for subsamples of managerial changes

	British		Age		Capped	
	Yes	No	>=50	<50	Yes	No
Rank Opponent	0.05***	0.06***	0.05***	0.04***	0.04***	0.05***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Home	0.59***	0.54***	0.61***	0.55***	0.57***	0.58***
	(0.04)	(0.09)	(0.06)	(0.05)	(0.05)	(0.05)
Manager change	0.17***	0.37***	0.21**	0.20***	0.23***	0.18**
	(0.06)	(0.09)	(0.09)	(0.06)	(0.08)	(0.07)
Counterfactual	0.25***	0.17*	0.21**	0.23***	0.15**	0.24***
manager change	(0.06)	(0.09)	(0.08)	(0.07)	(0.07)	(0.08)
F-test for equality	0.87	2.22	0.00	0.09	0.57	0.29
Observations	3,268	950	2,052	2,242	2,052	2,128
n-Seasons	86	25	54	59	54	56
n-Treatment-Group	48	13	29	32	30	31
n-Control-Group	38	12	25	27	24	25

Note: Team performance is measured by the number of points per match. Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All estimates include club-season fixed effects. 'Rank opponent' is the rank of the opponent in the preceding season. 'Home' indicates whether a match was played at home.

Finally, Table 2.4 presents the results for three subsamples that are based on the rank of the team in preceding year. The latter functions as a crude indicator of the quality and status of a club.[10] In columns one and two, we distinguish between clubs that were promoted in the previous season from the second tier of English football, the Championship, to the Premier League. Three teams were promoted in each season during the sample period, resulting in eight treatment groups to be considered, compared to 53 non-promoted teams. Extending the definition of promotion to one of the two preceding seasons, the number of treatment cases increases to 13, while 48 then belong to the non-promoted category. The results for these subsamples are presented in the third and fourth column. The last two columns provide results for subsamples where we distinguish between clubs that finished in the top half ($n=23$) and in the bottom half ($n=38$) of the Premier League table in the preceding season, treating promoted teams as part of the bottom. In contrast to the results in the Tables 2.2 and 2.3, we now find some insignificant values. The coefficient for the treatment group of the promoted teams in the preceding season (column 1) is positive, but insignificant, meaning that, for this subsample of cases, performances did not improve after the change in manager.

[10] The same remark about the number of treatment groups and the number of control groups made for Table 2.3 (footnote 8) applies to Table 2.4.

Interestingly, the coefficient for the control group is positive and significant, but the *F*-test for equality reveals that there is no significant difference between the treatment and control groups, which might have to do with the small number of observations in this subsample. The other insignificant results are found for the top half of the league table (column 5). Here, both coefficients for the treatment and control group are positive, but insignificant, strengthening the idea that, for this subset of club-season combinations, performances develop irrespective of the manager in charge. The *F*-test reveals no significant difference, which is also the case for all other subsamples that do contain positive and significant results.

Table 2.4: Results for subsamples of teams using all changes

	Promoted-1		Promoted-1-2		Rank-1	
	Yes	No	Yes	No	Top-10	Bottom-10
Rank Opponent	0.04***	0.05***	0.04***	0.05***	0.05***	0.05***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)
Home	0.64***	0.56***	0.68***	0.55***	0.54***	0.58***
	(0.07)	(0.04)	(0.06)	(0.04)	(0.07)	(0.04)
Manager change	0.19	0.21***	0.19*	0.21***	0.14	0.25***
	(0.12)	(0.06)	(0.10)	(0.06)	(0.10)	(0.06)
Counterfactual	0.27**	0.19***	0.21**	0.20***	0.12	0.24***
managerial change	(0.12)	(0.06)	(0.09)	(0.07)	(0.08)	(0.07)
F-test for equality	0.20	0.05	0.02	0.01	0.02	0.00
Observations	608	3,572	988	3,306	1,672	2,508
n-Seasons	16	94	26	87	44	66
n-Treatment-Group	8	53	13	48	23	38
n-Control-Group	8	41	13	39	21	28

Note: Team performance is measured by the number of points per match. Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All estimates include club-season fixed effects. 'Rank opponent' is the rank of the opponent in the preceding season. 'Home' indicates whether a match was played at home. 'Promoted-1' indicates a subsample of clubs that were promoted in the preceding season. 'Promoted-1-2' refers to a subsample of clubs that were promoted in one of the preceding two seasons. 'Rank-1' indicates a subsample of clubs that finished in the top half or bottom half in the preceding season, treating promoted clubs as bottom.

2.4 Case studies of managerial replacements

Our results in the previous section reveal that, *on average*, performances improve after the replacement of a manager, but the improvement is not causally related to the change. This is in line with previous studies. Nevertheless, there is a clear heterogeneity in the effects of a managerial change when we look at individual managerial changes. Figure 2.2 presents a scatterplot of all 61 changes included in our sample. The horizontal axis refers to the change in cumulative surprise after the managerial replacement. The vertical axis indicates the change in cumulative surprise for the control group.[11] For the sake of clarity, we added a diagonal that indicates equality of equal change in cumulative surprise for the treatment group and the control group. Observations above the line represent cases in which the control group did better than the treatment group, suggesting that the change was ineffective or even counterproductive. Observations below the line represent cases in which the managerial change was effective. Furthermore, the closer the observations are to the line, the more equal the development of the two groups is. Many observations are fairly close to the diagonal, which suggests that the managerial change was ineffective, thus supporting our *average* result. However, a substantial number of observations are at a fairly large distance from the diagonal, suggesting that some changes are quite effective, while others are counterproductive.

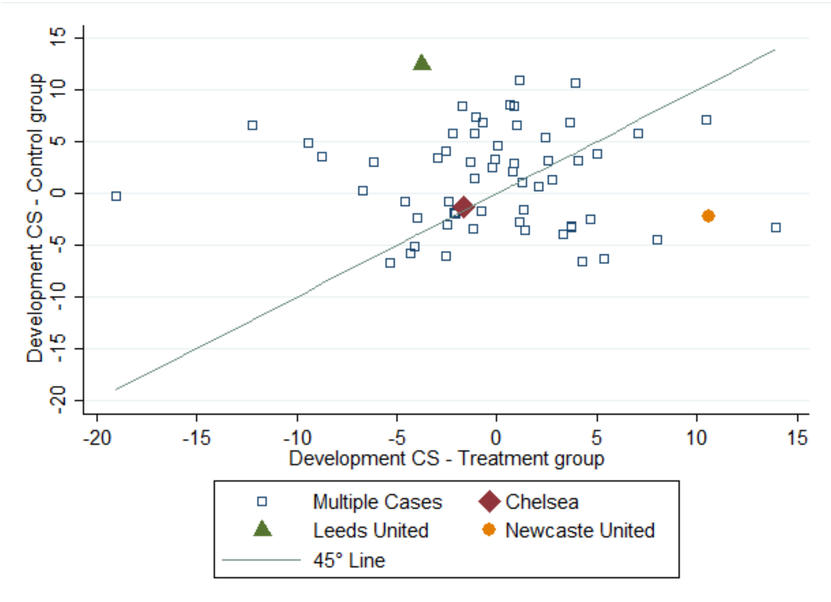
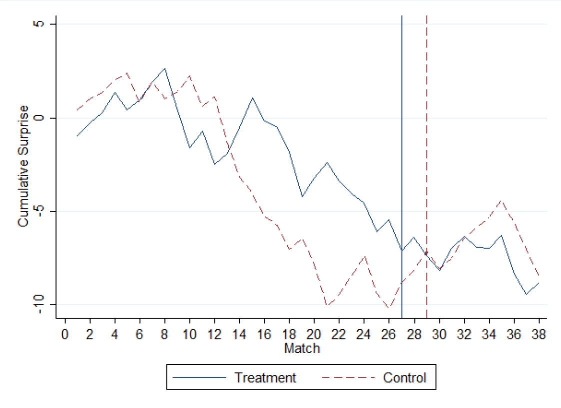


Figure 2.2: Development of Cumulative Surprise for all individual treatments and matched counterfactuals

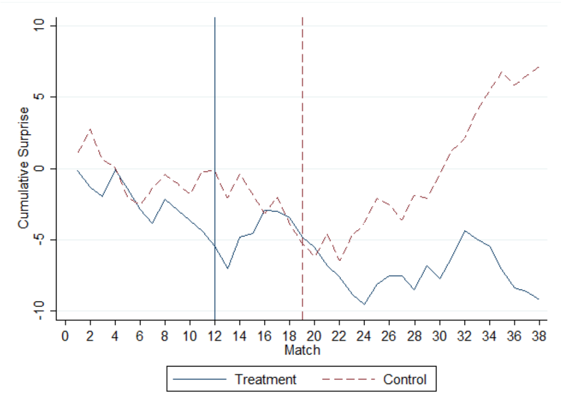
To investigate whether there are particular reasons for effectiveness or ineffectiveness of a managerial change, we selected three managerial replacements to discuss in more detail. First, we look at Chelsea FC, with treatment season 2011/12 and counterfactual 2010/11. This observation is indicated as a 'diamond' in Figure 2.2. The close proximity of the 'diamond' towards the diagonal line suggests hardly any effect at all. Second, we discuss Leeds United FC, with treatment season 2003/04 and counterfactual 2000/01. This observation is indicated as a 'triangle' in Figure 2.2. The position of the

[11] Since the cumulative surprise at the managerial change and the cumulative surprise at the counterfactual event does not exceed 0.5, we compare the change in cumulative surprise for both events. The values are calculated from Table A1 by taking FS-CS for the treatment group and MFS-MCS for the control group.

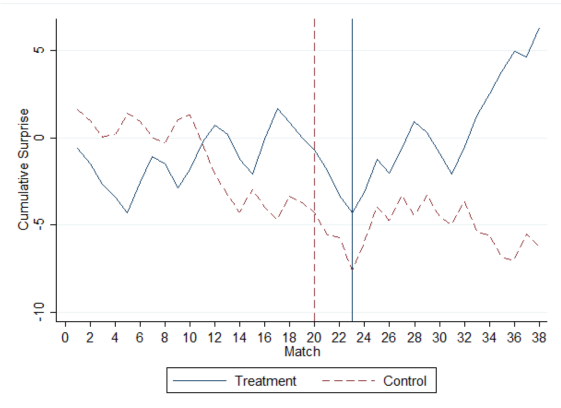
'triangle' suggests a strongly negative effect. Third, we examine Newcastle United FC, with treatment season 2005/06 and counterfactual 2012/13. This observation is indicated as a 'circle' in Figure 2.2. The position of the 'circle' suggests a substantially positive effect. All three cases concern the dismissal of the manager.



a: Chelsea FC, treatment season 11/12, control season 10/11



b: Leeds United FC, treatment season 03/04, control season 00/01



c: Newcastle United FC, treatment season 05/06, control season 12/13
Figure 2.3: Case studies

2.5 Chelsea FC[12]

André Villas-Boas moved from FC Porto to Chelsea FC in the summer of 2011. The Portuguese manager, only 33 years old at the time, had just guided the 'Dragões' (Dragons) to victory in the UEFA Europa League. Rumour has it that the London club paid a transfer fee of approximately 15 million euro. Villas-Boas soon presented a three-year plan to take the London club to the top of Europe. Yet, Chelsea-owner Roman Abramovich had already run out of patience after little more than eight months. The Russian club-owner held the manager responsible for the disappointing results. Thus, on 4 March, 2012, Chelsea FC sacked their Portuguese manager. Former Italian midfielder Roberto Di Matteo, previously an assistant to Villas-Boas, took over, initially only as a caretaker. At the end of the season, Chelsea were sixth in the table. However, Di Matteo guided them to their first ever victory in the UEFA Champions League (UCL). Moreover, Chelsea also won the FA Cup under his supervision.[13]

Carlo Ancelotti became the Chelsea FC manager in the summer of 2009. The former Italian midfielder had guided AC Milan to two UCL-victories (2003, 2007). In the 2009/10 season, he led Chelsea to the double, viz. both the EPL and the FA Cup. However, Chelsea lost both prizes in the next season. Abramovich sacked Ancelotti immediately posterior to the last match of the 2010/11 season. One month earlier, rivals Manchester United FC had eliminated Chelsea FC in the quarter finals of the UCL, a trophy then still absent in the club's boardroom. This has probably been a crucial element underlying this post-season sacking.

Figure 2.2 immediately makes clear that the difference between the control season (2010/11) and the treatment season (2011/12) is negligible. Moreover, the decline in cumulative surprise after the (hypothetical) change in manager is about equal for both seasons (see Figure 2.3a). The efforts that resulted in winning two trophies probably explain the disappointing results in the EPL in the treatment season, despite replacing the manager, who apparently was a mismatch. After all, the importance of the FA Cup may have decreased in the 21th century, but the UCL is, no doubt, the biggest prize in European club football.

2.6 Leeds United FC (LUFC)[14]

In the 2003/04 season, the debts of Leeds United FC were assessed as astronomically high, at around 100 million pound sterling. Consequently, LUFC had to go on selling quality players, weakening

---

[12] The information in this subsection stems from the articles concerning Chelsea FC seasons in Wikipedia.

[13] Chelsea FC sacked Roberto di Matteo on 21 November 2012, after Italian champions Juventus FC had eliminated them from the 2012/13 UCL.

[14] The information in this subsection stems from the articles concerning Leeds United FC seasons in Wikipedia.

their squad. The board sacked manager Peter Reid, a former England international midfielder, on 10 November 2003, a few months after his arrival at Elland Road. At that time, LUFC had gained no more than eight points from a dozen EPL matches. Eddy Gray, an all-time club-hero, took over as a caretaker. Initially, the results got better under his supervision: LUFC even moved out of the danger zone at the end of 2003. However, they subsequently lost seven matches in a row. Yet, the 'Whites' succeeded in bouncing back a little, one more time. However, in the end, relegation was inevitable.

David O'Leary was in charge at Elland Road from 1 October 1998, when he succeeded his former boss George Graham, until the summer of 2002. At that time, the board sacked him. O'Leary had been allowed to spend more than 100 million pound sterling in the transfer market, without winning any trophy. O'Leary's team seriously dipped during the 2000/01 season, but they recovered. These plunges may be ascribed to the lagged fatigue effects and leading anticipation effects of UCL matches, at least partly. In April 2001, LUFC reached the semi-finals of the UCL/European Champions' Cup for the first time since 1975.

Figure 2.2 makes clear that the difference between the control season (2000/01) and the treatment season (2003/04) is positive. Moreover, Figure 2.3b demonstrates that the cumulative surprise developed unfavourably after the managerial change in the 2003/04 season as compared to the same period in the control season.

2.7 Newcastle United FC (NUFC)[15]

Newcastle United FC experienced a turbulent summer in 2005. Rumours concerning the club-ownership, the departure of some star-players and the failure to qualify for Europe via the UEFA Intertoto cup (UIC) all contributed to the turmoil. Meanwhile, the Scottish manager Graeme Souness, a former Liverpool FC-hero, bought some first-class players, including England striker Michael Owen, who returned to England for 17 million pound sterling, after one season at Real Madrid.

Initially, Owen nicely co-operated with Alan Shearer, the latter in his final season as an active player. However, Owen got seriously injured on New Year's Eve. After that, the form of the team decreased severely. One month later, the NUFC board sacked Souness. A stiff battle against relegation then seemed to lie ahead for the 'Magpies'. The 2005/06 season then seemed to lack any prospect for the 'Magpies'. Glenn Roeder, director of the youth academy, took over as caretaker. He guided the team from the fifteenth place to the seventh place, thus even capturing an UIC spot. The team won no less

[15] The information in this subsection stems from the articles concerning Newcastle United FC seasons in Wikipedia.

than nine matches out of the remaining 14 matches in the EPL. Irish national goalkeeper Given and Shearer uttered afterwards that Souness had never been a fans' favourite and that his preference for certain players had been devastating for the team spirit. However, injuries had also been a crucial element in their dipping form.

In the 2012/13 season Alan Pardew guided NUFC to the 16th place. Thus, they avoided relegation. In the FA Cup and in the Football League Cup, they only lasted one round. However, NUFC did reach the quarter finals of the UEFA Europa League, which might explain their disappointing performance in the EPL and the domestic cup competitions, at least partly.

The chemistry between Souness and part of the team had apparently gone during the treatment season (2005/06). Moreover, the mighty fans of the 'Magpies' did not appreciate his work. Under such circumstances, the replacement of a manager may be an inevitable measure. During the control season (2012/13), NUFC were mediocre in all three domestic competitions. This may be explained from huge European efforts. Thus, it is hardly surprising that the difference between the treatment season (2005/06) and the control season (2012/13) is positive, as Figure 2.2 makes clear. Furthermore, Figure 2.3c demonstrates the cumulative surprise developed favourably after the managerial change in the 2005/06 season as compared to the same period in the control season.

2.8 Concluding remarks

In English premier league football managers are replaced for various reasons, but predominantly because of poor performance (Audas, Dobson and Goddard (1999), Dobson and Goddard (2011) Bachan, Reilly and Witt (2008), d'Addona and Kind (2014)).[16] In our paper, we investigate the effectiveness of in-season manager replacements, using data of 15 seasons from English Premier League football. When we compare the change in performance after managerial replacements with the change in performance of counterfactual replacements we find no difference. Although we find heterogeneity in the effects of managerial changes, the successfulness seems to be related to specific and highly unpredictable circumstances. This raises the question why coaches are dismissed anyway. There are several potential reasons for this. The first possible reason is that some club-owners are good in recognizing that a managerial replacement might be effective, while other club-owners are not. The second possible reason is misperception. As performance after a managerial change is often better than before, the perception is that this change was successful. True or not, club-owners are probably not interested in counterfactuals. A before-after comparison without considering a

[16] See Van Ours and Van Tuijl (2016) for determinants of coach replacement in other European football leagues.

counterfactual is misleading from a researchers' point of view, but not in the perception of club-owners, fans and mass-media. The third possible reason is asymmetry in the perception of the relationship between decision and result. Deciding for a replacement and not have an improvement in results is better than deciding not to act and not have an improvement in results. In the first case, club-owners have at least tried to improve the performance, in the second case they failed to act. The fourth possible reason is that dismissal is simply the destiny of a manager. The position of a manager has once been invented such that a manager gets the blame for disappointing results and not the club-owner (Carter (2006, 2007)).

Thus, managers seem to be sacked due to reasons outside of their influence, functioning as scapegoats. In management literature this is found to be the case after bad performances (e.g. Khanna and Poulsen, 1995) and might be an optimal strategy, together with the appointment of an outside successor, in the aftermath of wrongdoing (Gangloff, Connelly and Shook, 2016). In sports, scapegoating of managers is found as well (e.g. Bruinshoofd and Ter Weel, 2003). Consequently, their jobs are highly uncertain. The saying "you're only as good as your last match" seems to be typically true for professional football managers. Therefore, they will ask for some compensation in return for this uncertainty. However, many qualified managers are available, who are all willing to work in the EPL. This makes it rather easy for clubs to find a suitable replacement. Therefore, one might expect marginal demands from their side. Although CEO-compensation is based on multiple factors, such as ability (Chang, Dasgupta and Hilary, 2010), the opposite seems to be true, as salaries seem to be sky-high, probably including a scapegoat premium as found by Ward *et al.* (2011) for the CEO of listed companies as well as for college American football coaches. We have found that performances develop irrespective of the manager in charge, which is in line with the doubts of Kuper and Szymanski (2010) about the influence of football manager. Apparently, extremely high salaries reflect the compensation for job uncertainty rather than the compensation for superior quality.

Appendix A: Details on the data

Table A1 provides an overview of the 61 valid matched observations. Besides information on the seasons, match rank-numbers and managers, the columns CS and MCS report the matched values of cumulative surprise which, by definition, do not differ more than 0.5. Interestingly, we do observe some (large) differences when comparing the final two columns that report the cumulative surprises at the end of the season. A difference between these two values might indicate a positive (or negative) development of performances after the replacement of the manager compared to the counterfactual. We would like to note that we compare surprises based on expectations obtained from bookmaker odds. If these odds are heavily based on recent in-season results, badly (well) performing teams are likely to face low (high) expectations, which would overestimate (underestimate) their performance in terms of surprise. Then, the cumulative surprise is probably not a good performance measure to evaluate the effectiveness of in-season coach changes, since clubs are only interested in the actual number of points obtained. We indeed use this as the main performance measure in our analyses. However, we question the focus on recent in-season performances by bookmakers, given the broad range of cumulative surprises. The cumulative surprise is a useful measure to compare performances between different clubs and seasons.

Table A1: Overview of manager changes and matched observations

Club	S	W	Manager	T	N	A	C	MS	MW	MManager	CS	MCS	FS	MFS
Aston Villa	01/02	23	Gregory, J.	Q	B	47	6	13/14	23	Lambert, P.	1.60	1.67	-2.50	-3.45
Aston Villa	10/11	5	MacDonald, K.	D	B	49	0	02/03	5	Taylor, G.	-0.81	-0.68	-1.59	-2.40
Aston Villa	14/15	25	Lambert, P.	D	B	45	40	06/07	26	O'Neill, M.	-2.39	-2.42	-1.56	0.44
Birmingham City	07/08	13	Bruce, S.	Q	B	46	0	10/11	13	McLeish, A.	-2.60	-2.70	-7.17	-3.45
Blackburn Rovers	04/05	4	Souness, G.	Q	B	51	54	09/10	5	Allardyce, S.	-2.31	-2.45	-1.48	5.97
Blackburn Rovers	08/09	17	Ince, P.	D	B	41	53	11/12	16	Kean, S.	-6.98	-6.85	-5.64	-8.45
Blackburn Rovers	10/11	17	Allardyce, S.	D	B	56	0	07/08	18	Hughes, M.	1.60	1.29	-0.95	5.36
Bolton Wanderers	07/08	9	Lee, S.	D	B	48	14	11/12	17	Coyle, O.	-6.11	-5.84	-7.26	-4.39
Bolton Wanderers	09/10	18	Megson, G.	D	B	50	0	02/03	19	Allardyce, S.	-2.71	-2.81	-2.65	1.75
Chelsea	00/01	5	Vialli, G.	D	C	36	59	01/02	16	Ranieri, C.	-2.52	-2.04	-2.59	1.32
Chelsea	07/08	6	Mourinho, J.	Q	C	44	0	13/14	4	Mourinho, J.	-1.01	-0.66	9.43	6.47
Chelsea	08/09	25	Scolari, F.	D	S	60	0	10/11	15	Ancelotti, C.	-3.78	-4.03	4.22	-8.48
Chelsea	11/12	27	Villas-Boas, A.	D	C	34	0	10/11	29	Ancelotti, C.	-7.13	-7.14	-8.81	-8.48
Chelsea	12/13	12	Di Matteo, R.	D	C	42	34	06/07	12	Mourinho, J.	1.89	1.73	4.63	3.07
Crystal Palace	13/14	8	Holloway, I.	Q	B	50	0	04/05	9	Dowie, I.	-4.06	-4.07	9.86	-7.32
Derby County	01/02	7	Smith, J.	Q	B	60	0	00/01	7	Smith, J.	-2.75	-3.18	-8.88	-0.13
Derby County	07/08	14	Davies, B.	Q	B	43	0	00/01	12	Smith, J.	-6.89	-6.69	-19.13	-0.13
Everton	01/02	29	Smith, W.	D	B	54	0	00/01	30	Smith, W.	-4.57	-4.76	-2.51	-4.04
Fulham	02/03	33	Tigana, J.	D	C	47	52	04/05	34	Coleman, C.	-4.40	-4.30	0.62	-0.45
Fulham	06/07	33	Coleman, C.	D	B	36	32	04/05	31	Coleman, C.	-3.39	-3.48	-4.72	-0.45
Fulham	07/08	17	Sanchez, L.	D	B	48	3	10/11	18	Hughes, M.	-6.16	-6.59	-5.52	1.94
Fulham	13/14	13	Jol, M.	D	C	57	3	10/11	15	Hughes, M.	-3.78	-3.82	-5.98	1.94
Hull City	09/10	33	Brown, P.	D	B	50	0	14/15	28	Bruce, S.	-3.48	-3.94	-5.58	-5.94
Leeds United	03/04	12	Reid, P.	D	B	47	13	00/01	19	O'Leary, D.	-5.46	-5.26	-9.22	7.20
Leicester City	01/02	8	Taylor, P.	D	B	48	4	03/04	9	Adams, M.	-3.90	-3.90	-7.91	-6.28
Liverpool	10/11	20	Hodgson, R.	Q	B	63	0	11/12	29	Dalglish, K.	-8.82	-8.53	-4.54	-15.17
Manchester City	04/05	29	Keegan, K.	Q	B	54	63	06/07	34	Pearce, S.	-0.64	-0.59	3.09	-3.85
Manchester City	09/10	17	Hughes, M.	D	B	46	72	06/07	16	Pearce, S.	-0.26	-0.34	1.14	-3.85
Manchester United	13/14	34	Moyes, D.	D	B	51	0	01/02	18	Ferguson, A.	-5.08	-5.39	-6.19	0.35
Middlesbrough	00/01	16	Robson, Bryan	D	B	43	90	08/09	31	Southgate, G.	-9.19	-8.92	-5.47	-12.05
Newcastle United	04/05	4	Robson, Bobby	D	B	71	20	03/04	4	Robson, Bobby	-4.46	-4.10	-11.18	-3.90
Newcastle United	05/06	23	Souness, G.	D	B	52	54	12/13	25	Pardew, A.	-4.31	-3.97	6.30	-6.25
Newcastle United	07/08	21	Allardyce, S.	D	B	53	0	03/04	20	Robson, Bobby	-3.57	-3.17	-6.01	-3.90
Newcastle United	10/11	16	Hughton, C.	D	B	52	53	03/04	16	Robson, Bobby	-0.67	-1.12	0.48	-3.90
Newcastle United	14/15	19	Pardew, A.	Q	B	53	0	02/03	17	Robson, Bobby	3.85	3.82	-5.57	8.74
Norwich City	13/14	33	Hughton, C.	D	B	55	53	04/05	18	Worthington, N.	-3.32	-3.73	-5.80	-6.82
Portsmouth	04/05	13	Redknapp, H.	Q	B	57	0	03/04	14	Redknapp, H.	-1.69	-1.65	-4.66	1.84
Portsmouth	05/06	13	Perrin, A.	D	C	49	0	03/04	24	Redknapp, H.	-4.61	-4.71	-3.62	1.84
Portsmouth	08/09	8	Redknapp, H.	Q	B	61	0	07/08	8	Redknapp, H.	3.48	3.34	-5.24	6.96
Portsmouth	09/10	13	Hart, P.	D	B	56	0	03/04	27	Redknapp, H.	-6.20	-6.58	-7.90	1.84
Reading	12/13	29	McDermott, B.	D	B	51	0	07/08	28	Coppell, S.	-5.62	-5.81	-7.77	-7.66
Southampton	00/01	29	Hoddle, G.	Q	B	43	53	02/03	25	Strachan, G.	8.78	8.83	7.61	5.35
Southampton	01/02	8	Gray, S.	D	B	41	0	02/03	6	Strachan, G.	-1.61	-1.53	2.04	5.35
Southampton	03/04	25	Strachan, G.	Q	B	47	50	13/14	20	Pochettino, M.	-0.92	-0.45	-1.16	2.08
Southampton	12/13	22	Adkins, N.	D	B	47	0	02/03	6	Strachan, G.	-1.21	-1.53	-1.91	5.35
Sunderland	02/03	9	Reid, P.	D	B	46	13	07/08	9	Keane, R.	-0.93	-1.23	-19.96	-1.45
Sunderland	08/09	15	Keane, R.	Q	B	37	68	01/02	14	Reid, P.	-1.88	-2.25	-7.25	-8.94
Sunderland	11/12	13	Bruce, S.	D	B	50	0	07/08	14	Keane, R.	-4.84	-4.62	-0.76	-1.45
Sunderland	12/13	31	O'Neill, M.	D	B	61	64	07/08	29	Keane, R.	-3.21	-3.61	-2.41	-1.45
Sunderland	13/14	5	Di Canio, P.	D	C	45	0	07/08	14	Keane, R.	-4.40	-4.62	-1.86	-1.45
Sunderland	14/15	29	Poyet, G.	D	S	47	26	01/02	30	Reid, P.	-5.01	-4.93	-1.73	-8.94
Tottenham Hotspur	00/01	29	Graham, G.	D	B	56	12	01/02	25	Hoddle, G.	-5.39	-4.92	-4.10	-3.89
Tottenham Hotspur	03/04	6	Hoddle, G.	D	B	45	53	06/07	7	Jol, M.	-3.64	-3.49	-4.68	3.89
Tottenham Hotspur	04/05	11	Santini, J.	Q	C	52	0	12/13	11	Villas-Boas, A.	-1.42	-1.23	-0.27	9.74
Tottenham Hotspur	13/14	16	Villas-Boas, A.	D	C	36	0	09/10	16	Redknapp, H.	0.19	0.31	7.25	6.13
West Bromwich Albion	04/05	10	Megson, G.	D	B	45	0	05/06	7	Robson, Bryan	-3.56	-3.51	-6.08	-9.54
West Bromwich Albion	10/11	25	Di Matteo, R.	D	C	40	34	02/03	25	Megson, G.	-2.91	-3.29	2.43	-9.66
West Bromwich Albion	13/14	16	Clarke, S.	D	B	50	6	05/06	15	Robson, Bryan	-3.62	-3.70	-7.96	-9.54
West Bromwich Albion	14/15	19	Irvine, A.	D	B	56	0	08/09	19	Mowbray, T.	-4.27	-4.24	0.41	-6.67
West Ham United	06/07	17	Pardew, A.	D	B	45	0	13/14	19	Allardyce, S.	-6.91	-6.89	-4.52	-1.45
Wigan Athletic	07/08	12	Hutchings, C.	D	B	50	0	11/12	13	Martinez, R.	-4.37	-4.66	-0.48	6.07

Note: S indicates 'Season', W denotes the last match of the coach, T refers to the type of change with Q being a quit and D being a dismissal, N points at nationality with B for British, C for Continental (Europe) and S for South-America, C indicates the number of caps as a player, A refers to the age in years, CS points at cumulative surprise, FS indicates the final surprise (at the end of the season). The 'M' in the name of the column denotes the values belong to the matched observation.

Table A1 also provides information on some characteristics of the replaced manager, i.e. his nationality (column N), in particular, whether he has a British nationality, his age (column A) at the moment of replacement and the number of caps as a player (column C). We define ‘British’ managers as managers from the United Kingdom and from the Republic of Ireland, thus making a sharp distinction between these two countries and the rest of the world. Finally, column T reports whether the change was a dismissal or a quit.

Table A2 shows the 23 managerial changes that we have excluded from the analysis.

Table A2: Overview of manager changes not included in the analysis

Club	S	W	Coach	T	N	A	C	CS	FS
Bolton Wanderers	06/07	36	Allardyce, S.	Q	B	52	0	6.11	4.25
Bradford City	00/01	12	Hutchings, C.	D	B	43	0	-4.29	-9.04
Burnley	09/10	20	Coyle, O.	Q	B	43	1	0.67	-5.27
Cardiff City	13/14	18	Mackay, M.	D	B	41	5	-0.98	-7.47
Charlton Athletic	06/07	12	Dowie, I.	D	B	41	59	-4.50	-6.40
Crystal Palace	14/15	2	Millen, K.	D	B	47	0	-1.82	6.28
Leeds United	02/03	30	Venables, T.	Q	B	60	2	-9.16	-6.36
Manchester City	12/13	36	Mancini, R.	D	C	48	36	3.06	1.73
Newcastle United	06/07	37	Roeder, G.	Q	B	51	0	-4.03	-4.39
Newcastle United	08/09	3	Keegan, K.	Q	B	57	63	1.13	-10.69
Queens Park Rangers	11/12	20	Warnock, N.	D	B	63	0	-4.03	-3.55
Queens Park Rangers	12/13	12	Hughes, M.	D	B	49	72	-9.75	-16.31
Queens Park Rangers	14/15	23	Redknapp, H.	Q	B	67	0	-3.82	-6.13
Southampton	04/05	2	Sturrock, P.	D	B	47	20	0.52	-10.01
Sunderland	05/06	28	McCarthy, M.	D	B	47	57	-14.86	-17.26
Swansea City	13/14	24	Laudrup, M.	D	C	49	104	-7.04	-7.77
Tottenham Hotspur	07/08	10	Jol, M.	D	C	51	3	-6.89	-10.98
Tottenham Hotspur	08/09	8	Ramos, J.	D	C	54	0	-10.60	-4.09
West Ham United	00/01	37	Redknapp, H.	Q	B	54	0	-7.97	-8.99
West Ham United	02/03	35	Roeder, G.	D	B	47	0	-7.34	-4.49
West Ham United	08/09	3	Curbishley, A.	Q	B	50	0	1.64	5.67
West Ham United	10/11	37	Grant, A.	D	C	56	0	-6,94	-8,58
Wolverhampton Wanderers	11/12	25	McCarthy, M.	D	B	53	57	-6,13	-12,76

Note: S indicates season, W refers to the last match of the coach, T denotes the type of change with Q being a quit and D being a dismissal, N points at nationality with B for British, C for Continental (Europe) and S for South-America (no observations), C indicates the number of caps as a player, A denotes the age in years, CS refers to cumulative surprise, FS points at the final surprise (at the end of the season).

Chapter 3

Outcome uncertainty, team quality and stadium attendance in Dutch professional football

(Joint work with Jan van Ours and Martin van Tuijl)

3.1 Introduction

Attendances at professional (team) sport events are a popular research area. A central topic is the relation between the uncertainty of the outcome of a contest and consumer demand. Rottenberg (1956) and Neale (1964) were the first to formulate the well-known uncertainty of outcome hypothesis (henceforth UOH). They suggest that attending a match is more attractive if the outcome is uncertain. This concept has been introduced for single matches, referred to as *match uncertainty*. However, two other types of outcome uncertainty are recognized in competitive sports as well (e.g., Cairns, Jennett and Sloane, 1986; Borland and Macdonald, 2003). First, *seasonal uncertainty* is the uncertainty related to some end-of-season outcome, such as winning a league, promotion or relegation. Second, *long-run uncertainty,* which refers to the (lack of) dominance of certain teams during a considerable number of seasons. The UOH is often related to the concept of competitive balance, as proposed by Rottenberg (1956). No universally accepted definition of this concept exists. Consequently, competitive balance is measured in different ways (Owen, 2013). This concept generally relates to the degree in which competitors (such as sports teams) are balanced in terms of resources, quality, and talent etc. The *ex-ante* outcome of a match or competition between fairly equal competitors is more uncertain than the outcome of a contest between rather unequal competitors. If consumers, i.e. sports fans, derive utility from outcome uncertainty, a more balanced competition will attract more attendants. Therefore, sports bodies have an incentive to increase competitive balance, as they want to attract attendants in order to serve their members. Rules and regulations, such as salary-caps/wag-bill caps and talent allocation schemes (drafts), which are fairly common for team sports in the US, may be used to achieve this goal. In Europe, sports bodies are more reluctant to apply such restrictive regulations, since they are frequently bound by both domestic and European labour legislation.[1]

[1] At least in European football, the national football associations may punish clubs for financial disorder, e.g. by deducting points. Not too long ago, the UEFA (Union of European Football Associations), the governing body in European football that organizes the lucrative European Champions League and Europa League, has introduced the Financial Fair Play system. In short, this system sets limits to the financial losses of clubs, constraining the spending on salaries and transfer fees, by threatening with exclusion from UEFA competitions. Although one could doubt the efficacy of this system, as

Despite a large body of research (see Section 3.2), it is still not clear whether outcome uncertainty matters for attendance in professional football. Pawlowski (2013) provides an overview of studies concerning football attendance that shows mixed evidence with respect to the UOH. His review is dominated by studies that focus on match uncertainty and seasonal uncertainty. Schreyer, Schmidt and Torgler (2016) give an overview of studies on match uncertainty only. Few studies cover long-run uncertainty. Studies generally use match-level data from a limited number of seasons in a single country, and investigate whether stadium attendance depends on the uncertainty of outcomes.[2] Several indicators of uncertainty have been used and have found to be statistically significant (Borland and Macdonald, 2003; Pawlowski, 2013; Schreyer, Schmidt and Torgler, 2016). Nevertheless, the debate with regard to the UOH is still going on.

Recently, scholars interested in attendance demand looked for theoretical foundations for the hypothesis that are based on behavioural economic principles and decision making under uncertainty. Budzinski and Pawlowski (2017) provide a review on this recent stream of literature. The first to include behavioural economic theory are Coates and Humphreys (2012), who discuss the role of loss aversion as follows from prospect theory (Kahnemann and Tversky, 1979) for attendance demand in the context of the National Hockey League (NHL). Furthermore, Coates, Humphreys and Zhou (2014) develop a model of attendance demand that includes loss aversion combined with reference-depended preferences as described by Koszegi and Rabin (2006). They find that attendance is a function of the home win probability and its squared value. In their model, a concave relation between the home win probability and attendance suggests the classical UOH. It emerges as a special case within the model, where fans prefer tighter matches above certain home wins. For this to happen, the marginal utility of attending an unexpected win has to be at least as big as the marginal utility of an unexpected loss. A convex relation suggests that fans are loss averse. In that case, fans value home wins and the potential to attend an upset, i.e. a home win in case the home team is expected to lose. Fans attend such upsets if the expected utility of this unlikely event outweighs the utility of attending a home loss in a relative uncertain match. With controls for several match and team characteristics, such as team quality, an empirical test with data from the Major League Baseball (MLB) suggests a convex relation and, thus, the rejection of the UOH (Coates, Humphreys and Zhou, 2014). Humphreys and Zhou (2015) extend this model with a league standing effect, i.e. with seasonal uncertainty.

---

big clubs still spend large amounts of money, it can be seen as a form of regulation intending to increase competitive balance.

[2] Over time, an increased number of studies looked at the demand for TV audience (e.g. Forrest, Simmons and Buraimo, 2005; Buraimo and Simmons, 2015; Cox, 2015). In general, such data is not publicly available and, therefore, rather difficult to obtain. In the present study, we focus on stadium attendance.

Furthermore, they argue that a convex relation between the home win probability and attendance does not rule out the existence of fan preferences for uncertain matches. It only means that loss aversion dominates the preference for uncertainty.[3] They use data from the MLB to test their model. Since it does not prescribe how the league standing effect should be measured, a proxy is used.

In our study, we follow a similar approach and investigate the UOH in multiple ways with data from the highest level of Dutch professional football, covering the seasons 2000/01 - 2015/16. We use both, a measure of match uncertainty and one concerning seasonal uncertainty. For match uncertainty, we use the Theil-index and a measure initially proposed by Forrest, Simmons and Buraimo (2005), which consists of a points-per-game spread corrected for home advantage. These measures are traditionally used in attendance demand studies. Furthermore, we also empirically test the consumer choice models by Coates, Humphreys and Zhou (2014) and of Humphreys and Zhou (2015) with the home win probability and its squared value. As an alternative for the home win probability, we introduce the expected number of points for the home team. This the expected number of points is based on bookmaker odds and takes into account the possibility for a draw, which is the result in about one out of four matches in Dutch football.[4] We do not formally write down a model, but simply include it in the econometric specification, where it turns out be an effective indicator. Seasonal uncertainty is measured following Jennett (1984), who uses the significance of a match in relation to some end-of-season outcome. We adapt Jennett's significance measure to use it for winning a league as well as for other end-of-season outcomes, including qualifications for - the end-of-season play-offs for - European football competitions or relegation matches.[5]

In general, our results provide evidence for the rejection of the UOH related to match uncertainty. We find a convex specification for both the home win probability as well as our match-expectation variable. This suggests that fans exhibit reference-dependent preferences and loss aversion. Furthermore, it follows that team performance and team quality are important for the determination of stadium attendance. In general, the marginal effects are small. However, when measured in relation to the variable part of attendance, some effects tend to be quite substantial. For seasonal uncertainty,

[3] Humphreys and Zhou (2015) show that the structural model contains three parameters of interest. One concerns the home win preference, one regards the preference for outcome uncertainty and regards loss aversion. However, the reduced form of the model only contains two parameters that can be estimated. The first coefficient regards the home win probability, while the second is the squared value of the home win probability. Combined, these two parameters allow for the identification of the home win preference, which one finds if the sum of the coefficients is more than zero. However, it is impossible to identify the preference for outcome uncertainty and loss aversion separately. One can only observe which of these effects dominates.

[4] Measuring the expected number of points in this way has been used as an expectations-based reference point in football (Bartling, Brandes and Schunk, 2015) but, as to the best of our knowledge, it has never been used before in attendance demand studies.

[5] Section 3.4 and Appendix B include a detailed description of all measures.

many results are in line with the UOH. In particular, this holds true for significance related to relegation and the qualification for European football. We also investigate whether the introduction of a large play-off scheme in Dutch professional football in the season 2005/06 has had an effect on attendance demand during regular league matches.[6] In line with the UOH, one would expect a positive effect on attendance demand (see Bojke, 2007). We find a significant, though fairly small effect. When evaluating the overall effectiveness of the play-offs, we discuss whether this positive effect offsets some potential drawbacks of the system. Finally, we investigate between-season variation in attendance, focusing on the relation between changes in stadium capacity and average attendance.

Our paper adds to the existing literature in multiple ways. First, we use data from a large sample of seasons, that allows us to test the UOH for within-season variation in attendance using different indicators over a longer period. This seems important, as some seasons appear to be more uncertain than other seasons. Furthermore, certain seasons might contain more uncertain matches than other seasons, which partly results from the random order of play. Second, we introduce the expected number of points as a new measure for match uncertainty. This measure includes the probability of a draw, which is important in football, and appears to outperform other measures. It also follows that the model specification is important with regard to the optimum of the convex relation. Furthermore, we use a new way to measure the quality of teams. Bookmaker odds are used to calculate the expected number of points in the previous 34 matches. This measure of quality contains expectations and outperforms measures that combine pre-match rankings. In general, our results reveal that team characteristics are more important for the determination of stadium attendance than behavioural economic explanations regarding the outcome of the match. This contradicts wit previous findings. Finally, with respect to stadium attendance, we evaluate the introduction of the play-off scheme in the season 2005/06 empirically, contributing to the discussion on how competitions in professional football ought to be organized.

The structure of our paper is as follows. Section 3.2 provides a review of the literature concerning outcome uncertainty and attendance demand, distinguishing between match uncertainty and seasonal uncertainty. Next, section 3.3 provides some background information on Dutch professional football. Subsequently, section 3.4 deals with the data used and the research methods applied. Next, section 3.5 discusses our parameter estimates. Finally, section 3.6 concludes.

[6] These end-of-season play-offs typically involve four clubs. The ‘winner’ qualifies, while without play-offs, the best-ranked team would have qualified.

3.2 Review of literature

In this section, we discuss previous studies on the relation between outcome uncertainty in sports and attendance demand, with a focus on stadium attendance in professional football. Thereby, we ignore long-run uncertainty; we separately discuss match uncertainty and seasonal uncertainty in two subsections (Appendix C provides a summary overview).

3.2.1 Match uncertainty

The UOH was initially formulated with respect to match-specific outcome uncertainty. In theory, the concept is rather straightforward. However, empirical research has turned out to take many different shapes, due to the lack of a clear measure of match uncertainty. Initially, the focus was on the within-season (difference in the) league rankings of the opponents just prior to a match (e.g., Hart, Hutton and Sharot, 1975; Baimbridge, Cameron and Dawson, 1996; Garcia and Rodriguez, 2002). Alternatively, a measure was obtained from in-season performance, such as the difference in (average) points (e.g., Wilson and Sim, 1995) or goals (e.g., Falter and Perignon 2000). Notably, the importance of home advantage in football triggered Forrest, Simmons and Buraimo (2005) to come up with a measure that combines home advantage and the difference concerning the in-season number of points per game of the home team and the away team. Their reasoning is that the home advantage implies, that the home team is expected to win a match between teams of equal strength. Uncertainty of outcome increases with the relative strength (weakness) of the away (home) team. This point-per-game measure is calculated as follows:

$$PPG_{ijk} = |HA_k + PPG_{ik} - PPG_{jk}| \quad , \qquad (1)$$

in which i denotes the home team, j indicates the away team and k refers to the season. The points-per-game match uncertainty measure PPG_{ijk} is the absolute value of the home advantage (HA_k) plus the number of points per game of the home team (PPG_{ik}), minus the number of point per game of the away team (PPG_{jk}). The home advantage is measured as the difference between the number of points per game won at home and away in the previous season. Match outcomes of the current season are used to calculate points per game values of the home team and the away team.

Suppose that football fans decide to attend a match based on the uncertainty of the outcome. Moreover, let us assume that they use league tables to determine this uncertainty. In that case, these measures capture relevant match uncertainty, at least partly. However, these proxy variables have drawbacks. First, they might measure other aspects as well, such as the quality of teams.[7] Second, they do not capture forward-looking factors that influence perceived match uncertainty, such as injuries or suspensions of (key) players.

Peel and Thomas (1988) have introduced the posted fixed betting odds as input for measures of match uncertainty. The odds set by bookmakers, reflecting probabilities of a home win, of a draw and of an away win, provide easy to use information with regard to the evenness of a match. The more equal these probabilities are, the more uncertain the outcome of a match is. If the betting market is efficient, the bookmaker odds include all relevant information for the formation of expectations regarding the final result. Therefore, the uncorrected posted odds are generally acknowledged to be a useful input for match uncertainty.[8]

Again, numerous ways have been proposed and tested. At first, Peel and Thomas (1988) only use the home win probability. Peel and Thomas (1992, 1996) use the Theil-index, incorporating all three odds, responding to criticism that they ignored two possible outcomes.[9] This Theil-index has become the dominant measure of match uncertainty when using bookmaker data. It is calculated as follows:

$$\text{Theil-index} = \sum_{i=1}^{3} p_i \ln\left(\frac{1}{p_i}\right), \tag{2}$$

with p_i being either the probability of a home win, a draw, or an away win. With a high probability for one of the outcomes, the index takes on a value close to zero. In case of equal probabilities, the index is equal to ln(3). In general, the more uncertain the outcome of a match is, the higher the value of this index.

Recently, and in particular after the publications of Coates and Humphreys (2012), Coates, Humphreys and Zhou (2014) and Humphreys and Zhou (2015), scholars focus on the home win probability and its squared value in attendance demand studies. As explained in the introduction, the use of these values follows from a model with reference-dependent preferences and allows for loss

[7] Pawlowski (2013), for example, argues that this is the case for separate measures of pre-match league ranks.

[8] One should note that Forrest and Simmons (2002) present evidence of systemic biases in the bookmaker odds. They show that these biases matter when the odds are used for match uncertainty measures in an attendance demand model. Internet betting has probably increased competition between betting agencies and it has improved the transparency of the betting market. Both aspects may have had a positive influence on the efficiency of the market and a corresponding reduction of biases. Improved efficiency is also reflected in reduced bookmaker margins. Our in-sample average margin gradually declined from more than 16 percent in the beginning of the 21st century down to about eight percent in recent seasons. Thus, we assume that the bookmaker odds are set efficiently and reflect useable probabilities for the formation of match expectations.

[9] Some alternatives are considered in the literature. For example, Forrest, Beaumont, Goddard and Simmons (2005) use the ratio of the home-win probability over the away-win probability.

aversion. Within this model, a concave relation between the home win probability and attendance demand suggest support for the UOH. Alternatively, a convex relation suggests that loss aversion dominates the preference for uncertain outcomes. A home win preference is found if the sum of the coefficient for the home win probability and the coefficient for the squared value is positive (Humphreys and Zhou, 2015).

Most previous studies on match uncertainty use data from a single country and a limited number of seasons. The dependent variable is the logarithm of match-day attendance (stadium or televised). Furthermore, many studies test for a set of match uncertainty measures. They also frequently include some form of seasonal uncertainty. In general, the findings with respect to the UOH are mixed. Table C1 in the appendix provides more detailed information on these previous studies. We briefly discuss these studies here.

Falter and Perignon (2000) find support for the UOH, using the goal-average differential between opponents. Forrest and Simmons (2002) also find support for the hypothesis, using the ratio of the home win and the away win probabilities and the squared value of this ratio. Furthermore, Forrest, Simmons and Buraimo (2005) find support, for television audience, using the PPG indicator. Forrest *et al.* (2005) find some support for the UOH, using the ratio of the home win and away win probabilities, but not when applying separate point-per-game measures for the home team and the away team. Bojke (2007) finds the Theil-index to be significant, suggesting support for the UOH. Benz, Brandes and Franck (2009) use five different indicators of match-uncertainty. They find support for the UOH, using the difference in league rankings, the PPG indicator and the home team win probability, but only for high demand matches. They find no support using the Theil-index or a relative win-probability that does not incorporate the probability of a draw. Serrano *et al.* (2015) also only find support for high-demand matches, using both the Theil-index and its squared value. Buraimo and Simmons (2015) find support, only for the first two seasons of their sample that covers the period 2000/01 – 2007/08 for television-audience, using the difference in a teams' win-probabilities and the Theil-index. They do not find support for a measure of combined values for the in-season points-per-game of both teams. Based on a totally different approach, using survey data, Pawlowski (2013) finds that outcome uncertainty matters for fans, but increasing competitive balance would not result in increased demand for stadium attendance. Finally, Schreyer, Schmidt and Torgler (2016) find support for the UOH for the Theil-index and several other measures, but not for the home win probability. Results for this latter variable are more in line with reference-dependent preferences and loss aversion.

Notable studies that do not find support for the UOH include Wilson and Sim (1995), using the absolute point-difference (also squared) and Baimbridge, Cameron and Dawson (1996), using the difference in league rank (also squared). Moreover, Peel and Thomas (1996) do not find support, using repeat fixtures in the Scottish football league, for the difference in team rank and the difference in the Theil-index. They find partial support for the difference in the level of home win probability and this squared value. Furthermore, Czarnitzki and Stadtmann (2002) use the home team win probability, for which they do not find support. Garcia and Rodriguez (2002) use the difference in rank (also squared) and a dummy that indicates if the home team and the away team are close in the league table (with a maximum of three positions above or five positions below in the league table prior to the match). For both variables, no support in favour of the UOH is found. Furthermore, using the PPG indicator, Forrest and Simmons (2006), for stadium attendance, and Buraimo (2008), for both stadium attendance and television audience, do not find support. Using the home-win-probability (also squared) and the Theil-index, Buraimo and Simmons (2008) do not find support for the UOH. Madalozzo and Villar (2009) use the difference in positions in the league table, without finding any support. Pawlowski and Anders (2012) do not find support for the UOH, using the Theil-index as well as a home-favourite dummy that indicates whether the home win probability is higher than the away win probability. Cox (2015) uses the home win probability in multiple ways and does not find support for the UOH for stadium attendance, but results indicated a preference for loss aversion. However, he finds some support for the UOH for TV spectators. Martins and Cró (2016) use the home win probability, for which they find results in favour of loss aversion and home win preferences. Furthermore, the Theil-index is negative and a home favourite dummy positive. Thus, not supporting the UOH. Finally, Pawlowski, Nalbantis and Coates (2017) use survey data and conclude that fans' perceived match uncertainty coincides with the way economists measure it, but also find that loss aversion dominates their preference for such match uncertainty.

3.2.2 Seasonal uncertainty

Seasonal uncertainty is conceptually similar to match uncertainty, as both deal with expectations of some final performance. This can be a match outcome or some end-of-season achievement. However, as Cairns, Jennett and Sloane (1986) point out, empirical investigation of the UOH related to seasonal uncertainty does not specifically focus on the uncertainty surrounding this end-of-season goal, but it measures the probability of achieving this goal. The question then arises, whether attendance reacts to the probability that a team is still able to become champions, to qualify for UEFA competitions, or to avoid relegation.

Jennett (1984) was the first to come up with a measure that captures this type of seasonal uncertainty.[10] He uses the term 'significance' to indicate the importance of a match for an end-of-season result. Moreover, Jennett (*op. cit.*) assumes that the number of points needed to obtain this result is known in advance, i.e. the final table of a season is fully known. Significance is then calculated as the reciprocal of the number of matches to be won in order to achieve this result. If a team drops out of contention, because it has become mathematically impossible to gain the necessary minimum number of points, significance is set to zero for all remaining matches. Significance is also zero after a team has reached the minimum number of points needed to obtain the desired result. This would be the case if, for example, a team still needs to play two matches after obtaining a sufficient number of points to avoid relegation. Formally, this reads:

$$s_{ijk} = \frac{1}{m_{jk} - n_{ijk}} \; if \; pp_{ijk} \geq pt_k > pc_{ijk}, \quad (3)$$
$$s_{ijk} = 0 \; if \; pp_{ijk} < pt_k,$$
$$s_{ijk} = 0 \; if \; pc_{ijk} \geq pt_k,$$

in which i denotes the match, j indicates the club and k refers to the season. Thus, the significance s_{ijk} is equal to the reciprocal of the total number of matches in the season (m_{jk}) minus the number of matches already played (n_{ijk}) prior to match i. This only holds if the potential number of points for a specific team (pp_{ijk}) is larger than or equal to the total number of points needed to obtain the desired result. This total is constant throughout the season and represented by pt_k, which should be larger than the current number of points (pc_{ijk}) prior to a match. If the potential number of points is smaller than the total number of points, then significance becomes zero. Initially, Jennett used this measure to indicate 'championship significance' as well as significance related to relegation. Interestingly, it can also be used to calculate significance for other types of end-of-season achievements, such as the qualification for (the end-of-season play-offs for) UEFA competitions. One simply needs to change the value for total points (pt_k), which is simplified by the assumption that the final league table is known in advance. This knowledge, in turn, is probably the main point of criticism on Jennett's significance measure (e.g., Cairns 1987). As shown in appendix B, the value of pt_k is rather constant for Dutch football throughout the sample period. Therefore, it seems to be a close approximation for expectations that are based on the results in previous seasons.

[10] Although he describes it as an alternative measure of match uncertainty.

After Jennett's publication, other measures have been proposed. Most of these contributions use some kind of time-dimension, with increased uncertainty towards the end of the season. Janssens and Késenne (1987) suggest a small change to Jennett's measure, also assuming the final table is known. Cairns (1987) uses dummy variables that indicate whether teams are still in contention. Thereby, contention follows from some (arbitrary) rules that suggest that a team is still in contention if it wins a certain number of matches, while the current leader only wins a smaller number of matches. Goddard and Asimakopoulos (2004) propose a similar method, which Buraimo and Simmons (2015) use in an attendance demand function. Baimbridge, Cameron and Dawson (1996) use dummies in case the competing teams are both in contention for championship victory (both teams are ranked in the top four) or involved in a relegation battle (both teams are ranked in the bottom four). Kuypers (1997) suggests three similar measures that use some relation between the number of points that a team is currently behind the leader and the number of remaining matches in the season. Bojke (2007) uses a different approach, when investigating the impact on regular league match-attendance of the end-of-season play-off games for promotion in the English First Division (now known as the Championship) during the 2000/01 season. Seasonal uncertainty is measured by the probability of promotion. This probability is calculated with match-specific bookmaker odds and simulations to obtain match-by-match values for seasonal uncertainty.

Thus, in previous studies, seasonal uncertainty is measured by match-significance related to − the probability to reach − some end-of-season target. Table C2 in the appendix provides more detailed information on these studies. Here we only briefly discussed them. Jennett (1984), using his significance indicator, finds support for the significance related to 'championship victory' of both the home and away team, but not for relegation. Slightly modifying this indicator, Janssens and Késenne (1987) only find support for 'championship victory' of the home team. Cairns (1987) finds support using championship contention dummies, based on a set of (arbitrary) rules. Furthermore, Dobson and Goddard (1992) using a modified version of Jennett's indicator, just like Wilson and Sim (1995) using Jennett's indicator, find support for the UOH related to 'championship victory' of the home team, but not for the away team. Baimbridge, Cameron and Dawson (1996) do not find any support, using dummies that indicate whether opponents are both in the top or both in the bottom of the league table. A match day trend results in partial support, though only a crude indicator of seasonal uncertainty. Baimbridge (1997) investigates the UOH during EURO 1996. He measures match significance by the mean of winning probabilities and finds support for the UOH, while a matchday trend only partially supports the hypothesis. Kuypers (1997) finds support for the hypothesis for championship significance, but not for relegation. He measures both by using his own indicators.

Falter and Perignon (2000) use dummies for the different seasons of the year, which they find to be in line with the UOH. Using the indicator of Janssens and Késenne, Czarnitzki and Stadtmann (2002) do not find support for the hypothesis. However, Garcia and Rodriguez (2002) do find support, using one of Kuypers' indicators. Forrest, Simmons and Buraimo (2005) find, with a set of dummies, support for the UOH using data on television audience. Bojke (2007) uses his own way of measuring promotion probabilities. He finds that an increase in promotion probabilities results in increased attendance during regular league matches. This is in line with the UOH. Madalozzo and Villar (2009) find support for the UOH, using dummies when a team has either the chance to become the league leader, or the chance to leave the relegation zone. Furthermore, the result for the match number during the season supports the UOH, but the chance to qualify for the Copa Libertadores de América does not. Pawlowski and Anders (2012), who measure significance by the indicator of Janssens and Késsene, find support for the UOH related to championship significance for both the home and the away team. However, they do not find support for UEFA Champions League significance for the home team, whereas they only find partial support for the away team. In their study on television audience, Buraimo and Simmons (2015) do not find support for the UOH using several contention dummies. Martins and Cró (2016) use the uncertainty measure proposed by Janssens and Késenne (1987) for both championship victory and the qualification for the Champions League. For both variables, they find support for the home team, but not for the away team. Furthermore, they use rank order changes, which are insignificant. Finally, similar to his results for match uncertainty, Pawlowski (2013) finds, in his study on survey data, that seasonal uncertainty matters for fans. However, he also finds that increasing competitive balance does not result in increased demand for stadium attendance.

3.3 Dutch professional football

In November 1954, Dutch football, as organized by the KNVB (Koninklijke Nederlandse Voetbal Bond, Royal Dutch Football Association), turned from amateurism into (semi-)professionalism, with 56 clubs in total. At the start of the 1955/56 season, only 36 clubs were still active at the highest level, distributed over two divisions. The uniform nation-wide first tier, the so-called *Eredivisie*, started in the 1956/57 season, with 18 clubs. Only Ajax (Amsterdam), Feyenoord (Rotterdam) and PSV (Philips Sport Vereniging, Philips Sports Club, Eindhoven) have always been operating at the highest level of Dutch professional football ever since.[11] During these six decades, Ajax have won the Dutch title

[11] One should note that DOS (Door Oefening Sterk, Strong By Training) from Utrecht were active at the highest level between 1956 and 1970. FC Utrecht, which resulted from a merger between DOS and two other professional football clubs from the city of Utrecht, viz. Elinkwijk and Velox, has been active in the *Eredivisie* ever since.

25 times, PSV 20 times and Feyenoord ten times. Thus, the only three Dutch clubs to win UEFA silverware have captured no less than 90 percent of the title victories. This points at their long-term dominance in Dutch professional football. Feyenoord seems to perform the role of sleeping giant every now and then, but over the past few decades, Ajax and PSV have only rarely dropped out of the top-three.

The format of the *Eredivisie* is a so-called double round-robin system. Thus, every club plays all of the 17 adversaries twice, once at home and once away. Clubs obtain three points for a victory, a single point for a draw, but nil points for a loss. Direct relegation into the second tier, the so-called *Eerste Divisie*, is the consequence of the bottom position. The clubs ranked 16th and 17th have to play promotion/relegation play-offs with six clubs from this *Eerste Divisie*. These play-offs were introduced in the season 1989/90, although in a format quite different from the current one. The play-off scheme for qualification for European football was introduced in the 2005/06 season, amongst others to attract extra attendance. After the introduction, many final rankings made a club qualify for some (end-of-season) competition. The champion always directly qualified for the UEFA Champions League. The effects of finishing at one of the positions between 2nd and 13th changed multiple times. These changes are partly the result of changes in the number of Dutch teams that were allowed to participate in European competitions, partly because of changes in the set-up of these European competitions (e.g. the UEFA Intertoto Cup was abandoned after the 2006/07 season) and partly as a result of the introduction of, and changes in the design of, play-offs by the KNVB.

Although Dutch football is not known for its hooliganism, there have been some issues with violent fans. Therefore, the authorities sometimes simply forbid the presence of fans of the away team. Notable examples include the classical matches between Ajax and archrivals Feyenoord. Combined with some other safety measures, such as improved gate controls and increased presence of the police, these restrictions have resulted in a less violent atmosphere in the stadiums during matches. This, in turn, has been a key factor in attracting more attendance ever since the beginning of the 1990's. In general, safety, security, accessibility and facilities during matches have improved substantially in the past few decades. Numerous clubs have renovated their stadium or have built a new one in which these things are easier to realize.

The stadiums are generally quite crowded during matches. As we shall see in the next section, occupancy rates are approximately 80-90 percent throughout the sample period. The majority of attendants consist of season tickets holders. For example, data reported by magazine *Voetbal International* for the seasons 2011/12 and 2012/13 suggests that about 70 percent of the available seats within the *Eredivisie* is sold to season ticket holders. Furthermore, a yearly survey among fans, conducted by the KNVB in the seasons 2013/14 – 2016/17, reveals that about 90 percent of the season ticket holders plans to renew their season ticket for the upcoming year. Although this information is not sufficient to draw any conclusions regarding the whole sample period, it indicates that season ticket holders are loyal and an important part of the fan base and stadium attendants. They generally buy their season ticket before the start of the competition for a pre-announced fee. Single-match tickets are generally sold in the weeks prior to a match. Although the prices for these tickets are not necessarily fixed during the season, they have to be announced some period (i.e. weeks) in advance of the match. Furthermore, it seems reasonable to assume that the public knows what pricing strategies are used by clubs and that these strategies are rather constant throughout a season, making the prices for entrance tickets predictable. In general, the majority of stadium attendance consists of season ticket holders and the single ticket attendants. However, two other groups of attendants can be identified. These are business club members (sponsors) and attendants with free tickets, mostly children from local amateur clubs and primary schools. Although these two groups contain different types of consumers than the season ticket holders and single ticket attendants, our dataset does not allow to make such a distinction. As long as this group is not too big, that should not be a problem.

Probably the closest alternative for attending a match in a stadium, is to watch the match live on television. As of the start of the 1996/97 season, *Sport 7* provides this option. However, already after several unsuccessful months, the channel decides to quite with live broadcasting football matches. Next, *Canal Plus* provides live matches during the seasons 1997/98 – 2004/05. As one of the few broadcasting canals for which someone in the Netherlands had to pay an additional fee, next to their basic monthly fee for a subscription for television, this was not a profitable project. *Canal Plus* decides to quit and *Versatel/Tele2* buys the broadcasting rights for the next three seasons. The company wants to earn money via separate subscriptions, but again, without any profit. After these three unsuccessful attempts, no organization is willing to buy the broadcasting rights. Therefore, the clubs in the *Eredivisie* decide to start their own television channel. As of the start of the 2008/09 season, *Eredivisie Live* broadcasts the live Dutch football matches. Besides monthly subscriptions, they also offer pay-per-view options. With an outside investment by media tycoon Rupert Murdoch, amongst others owner of *Fox*, the broadcasting rights were sold again. As of the season 2013/14 *Fox Sports Eredivisie* broadcasts live football. In general, someone can watch all matches for a monthly fee of 10-20 euros. However, with an average of approximately 263.000 spectators per match during the seasons 2010/11 – 2015/16, the interest of the public remains only modest.[12] In contrast, stadium

[12] We thank Ruud Koning for the provision of the data on the number of spectators.

attendance rates are high, which happens to be the case for the big teams as well as the small clubs. Thus, although live broadcasting of football has become easily available for most people, we suggest that it did not serve as a substitute for attending the stadium. Instead, we suspect that it works as a complement, i.e. people watch football on television, become enthusiastic about the atmosphere and decide they want to attend the stadium themselves. However, this is a bit of a speculation and could be a topic for future research. Additional information about Dutch professional football is provided in Van Ours and Van Tuijl (2016).

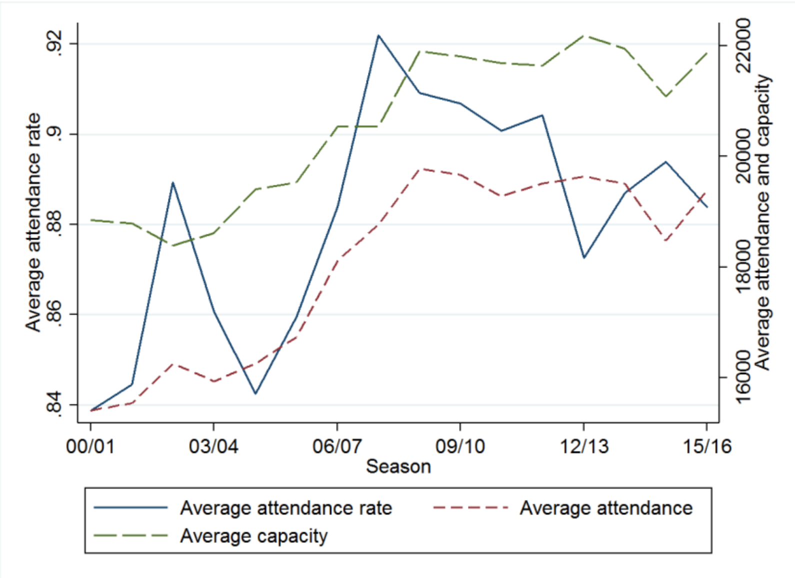


Figure 3.1: Average match attendance rate (left hand scale), average match attendance and average stadium capacity (right hand scale); seasons 2000/01 – 2015/16

3.4 Data and set-up of the analysis

We use match data from the highest tier of Dutch professional football for the seasons 2000/01 – 2015/16. For each match we have information on the attendance, general weather conditions during the day of the match, the competing teams, including the location (stadium and stadium capacity), city, province, as well as fixed betting odds (see Appendix B for details). In every season, at least one, but maybe two or even three teams are relegated, depending on the results of the end-of-season promotion/relegation play-offs, while a similar number of teams are promoted. In total, 28 different teams were active in the *Eredivisie* during the sample period.

Figure 3.1 shows the development of the average attendance rate, average match attendance and average stadium capacity during the sample period. Averages are calculated per season for all 18 clubs within the *Eredivisie*. It follows that the attendance rate is quite high in all seasons and moves around 0.88. Average attendance increased from approximately 15,500 in 2000/01 to 19,500 in

2015/16. In the same period, average stadium capacity increased from approximately 19,500 to almost 22,000. In general, these graphs develop in a similar way. Any difference in development is reflected in the attendance rate. Although the changes in this rate seem quite large, one should note that the scale only covers the range from 0.84 to 0.92. Furthermore, one should note that season-by-season changes in stadium capacity might result from actual changes in the capacity, because of renovations/expansions or the move to a new stadium, which generally happens in between seasons. Also, these changes may arise because of changes in the composition of clubs due to promotion and relegation. Certain clubs have bigger (smaller) stadiums than others. Correspondingly, these clubs might also have a bigger (smaller) fan base, which is reflected in the average attendance figures. In general, Figure 3.1 suggests that the interest in stadium attendance increased during the sample period, while the attendance rate remains high and stable.

We deal with the between-season variation in the second part of this study. The first and main part is focused on the within-season variation in attendance. We estimate a Tobit model with the attendance rate as the dependent variable. This rate has a maximum value of unity in the case that all seats are taken. However, our dataset may contain some measurement errors in attendance numbers. Furthermore, the formal stadium capacity may slightly differ, on a match-by-match basis, from the reported number as included in our dataset. In particular, because of specific policies and regulations regarding the fans of the away team. For example, in some matches, some seats might not be available, because club officials and local authorities want to have a clear separation of the away team fans from the home team fans. Both aspects may result in an under-identification of sell outs if we would use a value of one as upper limit. To correct for this, we assume the stadium to be full at 95 percent of the official capacity, i.e. we use an upper limit of 0.95 in our estimations. This results in about 40 percent of the observations being censored. The model is described as follows:[13]

$$AR_{cms} = \beta X_{cms} + \eta_{cs} + \varepsilon_{cms}, \qquad (4)$$

where AR refers to the attendance rate, c denotes the club, m indicates the match, while s refers to the season. Furthermore, X_{cms} represents a vector of club and match characteristics, β denotes a vector of parameter estimates, η_{cs} refer to club-season fixed effects that control for unobserved club-season factors, while ε_{cms} denotes the error term. Club-season fixed effects are included to control

[13] We have also performed Tobit tests with the natural logarithm of the attendance rate as the dependent variable, taking into account a comparable upper limit, but the results do not differ much. Since the model with the attendance rate as dependent variable is easier to interpret, we decide to use that in the rest of the paper.

for changes in unobserved factors with respect to different clubs over time. These include, amongst others, the number of season ticket holders, the pricing of tickets as well as changes over time in club-specific price-setting strategies, player transfers, club-specific penalties imposed by the UEFA or KNVB, season-specific European or domestic cup efforts, and changes in the fan base. Separate club and season fixed effects would capture only some of these unobserved elements. Such an approach would treat all clubs in a similar way with respect to seasonal circumstances, such as income and inflation figures, and treat all seasons in a similar way with respect to club-specific factors, such as price-setting strategies. Club-season fixed effects take account of the heterogeneity of fans of different clubs with respect to intra-season differences.

Note that, with club-season fixed effects, we explain club-specific within-season changes in the attendance rate. Since the attendance rate is the ratio of actual attendance and stadium capacity, any change in the attendance rate may result from changes in this capacity. Then, the within-season variation in the attendance rate would not reflect changes in attendance demand, while that is what we want to observe. However, stadium capacity is generally stable during the season. In only five out of 270 club-season combinations that are considered, we observe a substantial in-season change in capacity. Therefore, we think it is safe to assume that in-season changes in the attendance rate reflect changes in attendance demand.[14]

The vector X_{cms} consists of a variety of explanatory variables. We include dummies for *derbies*, capturing within-province rivalry, which are expected to attract more attendance. Previous studies use various methods to include information on the timing and conditions of a match day, such as month-dummies and the rank-number of the match in the season. We decide to control for matches that are played on weekdays, excluding national holidays, while considering Friday as part of the weekend. Furthermore, we control for the weather conditions on match day. The *weekday*-dummy captures short-term opportunity costs, as it is likely that it is more difficult to attend a match after a working-day. Weather conditions are measured by the average temperature during match day as well as the amount of precipitation. *Temperature* is typically high during the start of the season in August, then drops during the winter period, after which it rises again until the end of the season in May. Given this pattern, this variable not only captures short-term opportunity costs, but also some seasonal element. Although there is less *precipitation* in the final months of the season compared to the rest of

[14] The five cases with an in-season change in capacity are: RBC Roosendaal, who moved to a new stadium during the 2000/01 season; FC Groningen, who moved to a new stadium during the 2005/06 season; SC Heerenveen, who renovated the stadium and expanded the capacity during the seasons 2003/04 and 2004/05; Vitesse, who reduced the official stadium capacity during the 2009/10 season.

the year, the pattern is less evident, which suggests that this variable mainly captures short-term opportunity costs.

Most previous studies that use a measure of current performance or quality of a team, use the pre-match rank or the cumulative number of in-season points of both the home team and the away team. These measures are crude, since they do not account for club-specific and season-specific elements. For example, a difference in pre-match rank between the number three and the number six of the table is treated the same as the difference between the numbers twelve and fifteen. Furthermore, the measures do not incorporate expectations. However, it seems reasonable that fans take these expectations into account, since a certain total number of points might be experienced as good for some clubs, while not for others. Therefore, we prefer to use the cumulative surprise as a measure of performance (*Odd-Surprise*). This variable is equal to the in-season sum of the differences between the actual number of points obtained and the expected number of points based on bookmaker odds. Since expectations are considered, it allows for an easy comparison of performances between clubs and seasons.[15]

For team quality, we introduce a new measure, defined as the sum of the expected number of points in the previous 34 matches. Again, expectations are based on bookmaker odds. These odds reflect the relative strength or quality of the competing teams. They also include match-specific aspects such as home advantage, current form and player injuries that may influence the expected match result. Therefore, inference of the quality of a team, based on the expectations of a single match (or a small number of matches) would be wrong. For example, given home advantage, the quality of the home team would be biased upwards, while the quality of the away team would be biased downwards. To correct for these match specific factors that may bias the inferred quality, we consider the previous 34 matches of a team for the construction of the quality indicator.[16] In that way, the measure includes expectations that are based on all sorts of matches, such that match specific element cancel out against each other. For example, the number of home matches will be about equal to the number of away matches. Furthermore, the variable contains matches against strong teams as well as weak opponents. Although not all the impact of the match specific elements may disappear, the descriptive statistics

[15] A potential drawback of this measure is that it starts at zero prior to the first match of the season and then starts to develop. Match-by-match development can only take place with values that are theoretically bound by minus three (i.e. zero points while three were expected) and three (i.e. three points while zero were expected), while the cumulative value can be quite different. However, some robustness checks showed that leaving out a subset of matches, either at the start or at the end of the season, does not have an impact on the result for this variable. Furthermore, since this measure should capture the reaction of attendance on in-season performance, it can be argued that at the start of the season, there simply is not much to react upon, which is reflected in the value of the cumulative surprise.

[16] Note that a season contains 34 matches and, therefore seems to be a logical number. We also tested with the previous 17 matches, but results were very much the same.

suggest that the proposed indicator measures quality in line with what someone would expect based on league tables. For example, the average quality corresponds with an average number of points for an average league rank (see also Appendix B). Furthermore, the variable contains an arbitrary element, since we use the 34 previous matches without any weight factors, but it seems to work quite well in our attendance demand model. As with the cumulative surprise, it allows for a continuum of match-by-match differences and an easy comparison between clubs and seasons. A potential drawback might be that no previous seasons' bookmaker odds are available for promoted teams. We solve this problem by using the expected number of point of the team that relegated and was replaced by the promoted team. Furthermore, because of the *Quality Team* variable, we lose the 2000/01 season in our analyses, since we lack odds prior to this season.

In line with previous studies, we hypothesize that fans' decision to attend a match might depend on the opponents' quality, the opponents' performance or any other characteristics of the opponent. Therefore, we also construct *Odd-surprise Opponent* and *Quality Opponent*. Furthermore, we test our model with opponent fixed effects that capture any unobserved elements, such as reputation and brand strength (e.g. Czarnitzki and Stadtmann, 2002; Pawlowski and Anders, 2012).[17]

Our main focus will be on the outcome uncertainty variables. For match uncertainty, we use the *Theil-index* and the points-per-game (*PPG*) measure, as proposed by Forrest, Simmons and Buraimo (2005). Furthermore, we test the consumer choice models as proposed by Coates, Humphreys and Zhou (2014) and Humphreys and Zhou (2015) with the *Home win probability* and its squared value. As an alternative, we introduce a new measure of match uncertainty, i.e. the expected number of points for the home team, which we refer to as *Match-Expectation*. We also use the squared value. With this, we attempt to include, in an easy and straightforward way, the probability of a draw. Since approximately 25 percent of the matches in football ends in a draw, this seems important. We did not derive the *Match-Expectation* variable from a structural model as Coates, Humphreys and Zhou (2014) and Humphreys and Zhou (2015) did. In comparison to the *Home win probability*, *Match-Expectation* adds the probability of a draw, which is weighted by 1/3 of the probability for a home win, because teams earn three points for a victory and one for a draw. Note that the pairwise correlations between *Home win probability* and *Match-Expectation* are almost equal to unity (see

[17] Each league contains a few teams that are traditionally seen as the top teams. For example, because they generally have the best players and won the most trophies. In the Netherlands, these clubs are Ajax, Feyenoord and PSV. Matches with one of these top teams as opponent, may attract attendance irrespective of their recent performance or quality. Therefore, many previous studies use dummy variables to control for such teams. We do so as well with opponent fixed effects. In that way, we account for all unobserved characteristics of all the opponents. This approach assumes that these elements are constant throughout the sample period and are the same for fans of the different clubs.

Appendix B). Based on intuition and the fact that it is conceptually similar to the use of the home win probability, we suggest that it can be used to empirically test the predictions of the attendance demand models. We also test how results change in case both variables are included in one model.

With respect to seasonal uncertainty, we use Jennett's (1984) measure of significance. We test for seasonal significance related to championship victory, the qualification for the Champions League, the qualification for the UEFA Europa League and relegation. As argued before, the main drawback of this measure is the assumption that the final table is known in advance. In particular, it is assumed that the minimum number of points needed to achieve a certain end-of-season target, i.e. pt_k in equation 3, is known. This does obviously not hold true. However, this number of points is rather constant across seasons. This means that clubs and fans might form a fairly precise expectation of the value of pt_k based on previous results. Therefore, this assumption might not be problematic for our purpose.[18]

3.5 Parameter estimates

3.5.1 Baseline results

Our baseline parameter estimates are shown in Table 3.1, with match uncertainty measured by the *Home win probability* and *Match-Expectation*. Table D1 in Appendix D provides the results of similar models with *Theil* and *PPG*. In general, the findings are very much the same. The positive and significant coefficient for *Derby* suggests that matches against local rivals attract more attendance. Furthermore, people are less likely go to a stadium if the match is played on a *weekday* and if *precipitation* is higher. This suggests that opportunity costs play a role in people's decision to attend a match. The result for *Temperature* is less clear. The negative coefficient means that attendance is lower if the temperature is higher, which might reflect a seasonal element. The temperature reaches its highest value at the start of the season, drops towards the winter, but it goes up again in the final months, though it does not reach the level of August. None of these results are influenced by the inclusion of opponent fixed effects (models (3) and (6)). However, these fixed effects have an impact on the significance of the in-season performance measure for the opponent. *Odd-Surprise Opponent* is insignificant in models (2) and (5), but significant in models (3) and (6). This suggest that, after controlling for unobserved elements, such as the brand strength and the reputation of the away team, people seem to appreciate good performances of the opponent, with a positive effect on stadium attendance. We also find a positive effect for the in-season performance of the home team. *Odd-Surprise Team* is positive and highly significant in all the specifications. The effect for *Quality Team*

[18] Figure B1 in the appendix shows the relevant values of pt_k for all 16 seasons.

is also positive and generally significant in all specifications in Table 3.1, with comparable values for the coefficients. Significance levels are somewhat lower for the specifications in Table D1 in the appendix. The positive effect for *Quality Team* suggests that stadium attendance increases with the quality of the team. The club-season fixed effects already account for constant elements of team quality within a season, such as a seasonal budget and the composition of the squad. In addition, this result suggests that within-season changes in the inferred quality, positively effect within-season changes in the attendance rates. Note that this effect adds to a positive effect of in-season performances. Results are similar for the inferred quality of the opponent (*Quality Opponent*), i.e. better opponents attract more attendance. Not surprising, the coefficient drops after controlling for opponent fixed effects that capture certain quality related aspects.

Concerning the findings for the match uncertainty variables, the results for the *Theil-index* contradict with the UOH. For *PPG* we find some results in line with the hypothesis, but the coefficients are small and the difference in results for the three specifications is puzzling. In line with the recent stream of literature, we focus on the results for *Home win probability* and *Match-Expectation*. These results are shown in Table 3.1. In general, the results contradict with the UOH. For both the *Home win probability* and *Match-Expectation*, the relation with the attendance rate is convex, meaning that loss aversion dominates the preference for uncertain matches. The bottom rows of Table 3.1 give the home win probabilities at the minimum of the convex relation, where we assumed the probability of a draw to be 0.25 (about equal to the in-sample mean value) in order to translate the minimum for *Match-Expectation* into a home win probability. The minimum value for the HWP in model (1) is 0.80, while in model (4) 0.81 respectively. This is about equal to maximum value for the home win probability within the sample (i.e. 0.88). Thus, this suggests that almost all observations lie within the downward sloping part of the convex curve, meaning that an increase in the home win probability results in a decreased interest in stadium attendance. A result that seems to contradict with the concept of loss aversion, since that would predict fans to favour high home win probabilities and observations within the upward sloping part of the convex curve. Furthermore, Humphreys and Zhou (2015) show that one can identify the presence of a home win preference by the sum of the coefficients for the home win probability and its squared value (see also footnote 3). If this sum is positive, a home win preference exists. The results in models (1) and (4) reveal that this is not the case. Thus, we find a convex relation that fits within the attendance demand model with reference-dependent preferences and loss aversion, but without observing loss aversion or a home win preference. In their model, Coates, Humphreys and Zhou (2014) provide an alternative explanation, i.e. fans' interest in upsets. Fans may want to attend live matches, because they enjoy the possibility of an unexpected win. These cases are characterized by relatively low home win probabilities. Otherwise a home win is not unexpected. The interest in upsets may account for a downward sloping part of the convex curve. However, it cannot explain why the minimum value lies at around 0.80.

Table 3.1: Baseline parameter estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Derby	0.043***	0.043***	0.041***	0.042***	0.043***	0.041***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Weekday	-0.029***	-0.029***	-0.025***	-0.029***	-0.029***	-0.025***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Temperature	-0.009***	-0.009***	-0.008***	-0.009***	-0.009***	-0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Precipitation	-0.010***	-0.009***	-0.010***	-0.010***	-0.009***	-0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Odd-Surprise Team/100	0.408***	0.273***	0.239***	0.410***	0.280***	0.241***
	(0.076)	(0.079)	(0.078)	(0.076)	(0.079)	(0.078)
Odd-Surprise Opponent/100		0.045	0.092***		0.042	0.092**
		(0.038)	(0.036)		(0.038)	(0.036)
Quality Team/100	0.316**	0.288**	0.242*	0.319**	0.290**	0.243*
	(0.146)	(0.139)	(0.135)	(0.146)	(0.139)	(0.135)
Quality Opponent/100		0.238***	0.092**		0.229***	0.089**
		(0.036)	(0.036)		(0.037)	(0.036)
Home win probability	-0.568***	-0.298***	-0.038			
	(0.051)	(0.060)	(0.061)			
Home win probability^2	0.357***	0.290***	0.078			
	(0.052)	(0.050)	(0.052)			
Match-Expectation				-0.203***	-0.127***	-0.028
				(0.021)	(0.022)	(0.023)
Match-Expectation^2				0.038***	0.035***	0.012*
				(0.006)	(0.006)	(0.007)
Championship	0.004	0.013	0.008	0.005	0.013	0.008
	(0.042)	(0.042)	(0.041)	(0.042)	(0.042)	(0.041)
UEFA Champions League	0.050	0.052	0.068*	0.050	0.052	0.067*
	(0.042)	(0.040)	(0.039)	(0.042)	(0.040)	(0.039)
UEFA Europa League	0.092***	0.088***	0.089***	0.092***	0.088***	0.089***
	(0.026)	(0.025)	(0.023)	(0.026)	(0.025)	(0.023)
Relegation	0.182***	0.174***	0.163***	0.183***	0.175***	0.164***
	(0.043)	(0.042)	(0.041)	(0.043)	(0.042)	(0.041)
Opponent FE	No	No	Yes	No	No	Yes
Minimum HWP	0.80	0.51	0.24			
Minimum Exp. Points				2.67	1.80	1.13
HWP with Draw Prob.=0.25				0.81	0.52	0.31

Note: Tobit regression with attendance rate as dependent variable and with the upper limit set at 0.95. All estimates contain 4,586 observations (1,890 censored) and 270 club-season fixed effects. Robust standard errors in parentheses, clustered by club-season. *Odd-Surprise Team, Odd-Surprise Opponent, Quality Team* and *Quality Opponent* are divided by 100; HWP means home win probability, Exp. Points is the expected number of points. ****p<0.01, **p<0.05, *p<0.1.*

From the results in models (2) and (5) it follows that the inclusion of the opponent related variables *Odd-Surprise Opponent* and *Quality Opponent* reduce the minimum value for the HWP, without altering the sign and significance of the match uncertainty variables. The value of approximately 0.51 is close to the sample mean value for the home win probability (i.e. 0.46). Furthermore, it means that approximately half of the observations lies within the upward sloping part of the convex relation between the home win probability and the attendance rate. This seems to be more in line with the concept of loss aversion as proposed by Coates, Humphreys and Zhou (2014). However, the sum of the coefficients still suggest the absence of a home win preference. With opponent fixed effects (models (3) and (6) in Table 3.1), the minimum value for the HWP drops even further. However, most coefficients become insignificant. Only the coefficient for *Match-Expectation^2* is significant at a 10 percent level. Therefore, we conclude that the results for the match uncertainty measures are highly dependent on the inclusion of the opponent related controls. It seems that *Home win probability* and *Match-Expectation*, together with their squared values, captured all the opponent elements in models (1) and (4). Based on the results of models (3) and (6), we conclude that team characteristics are more important in the determination of stadium attendance, than behavioural economic explanations regarding the outcome of the match (i.e. a preference for outcome uncertainty, loss aversion and a home win preference). At least within our sample.

Table 3.2: Sensitivity analysis match expectation parameters

	(1)	(2)	(3)	(4)	(5)	(6)
Odd-Surprise Team/100	0.428***	0.300***	0.257***	0.360***	0.218**	0.185**
	(0.077)	(0.082)	(0.080)	(0.094)	(0.086)	(0.084)
Odd-Surprise Opponent/100		0.033	0.084**		0.093**	0.144***
		(0.039)	(0.036)		(0.038)	(0.039)
Quality Team/100	0.324**	0.295**	0.247*	0.627***	0.421***	0.370**
	(0.146)	(0.139)	(0.135)	(0.149)	(0.149)	(0.149)
Quality Opponent/100		0.211***	0.076**		0.306***	0.154***
		(0.039)	(0.038)		(0.043)	(0.038)
Home win probability	1.578***	0.741	0.626			
	(0.563)	(0.574)	(0.519)			
Home win probability^2	-1.370***	-0.806**	-0.586			
	(0.383)	(0.392)	(0.367)			
Match-Expectation	-0.795***	-0.435**	-0.274	-0.190***	-0.086**	0.007
	(0.187)	(0.195)	(0.178)	(0.040)	(0.040)	(0.042)
Match-Expectation^2	0.188***	0.130***	0.078*	0.035***	0.032***	0.009
	(0.048)	(0.048)	(0.047)	(0.013)	(0.011)	(0.011)
Opponent FE	No	No	Yes	No	No	Yes
Club-Season FE	Yes	Yes	Yes	No	No	No
Club FE	No	No	No	Yes	Yes	Yes
Season FE	No	No	No	Yes	Yes	Yes

Note: Tobit regression with attendance rate as dependent variable and with the upper limit set at 0.95. All estimates contain 4,586 observations (1,890 censored). Models contain all other variables included in our baseline models (not reported). Models (1)-(3) contain 270 club-season fixed effects; models (4)-(6) contain separate club (28) and season (15) fixed effects. Robust standard errors in parentheses, clustered by club-season (models (1)-(3)) and clustered by club (models (4)-(6)). *Odd-Surprise Team, Odd-Surprise Opponent, Quality Team* and *Quality Opponent* are divided by 100; ***$p<0.01$, **$p<0.05$, *$p<0.1$.*

Table 3.2 presents the results of a sensitivity analysis focusing on the relevant parameter estimates. We re-estimate our baseline models, but combine the *Home win probability* and *Match-expectation* variables, to test whether one of them is preferred. The results are shown in models (1)-(3). First, note that none of the results for the team or opponent variables changes. Furthermore, the results for the

Match-Expectation variable are comparable to those in Table 3.1. However, for *Home win probability* the sign as well as the significance of the coefficients changed. Therefore, we conclude that *Match-Expectation* is preferred and is used in the rest of this study. Although the results in Table 3.1 are very much the same, it, thus, seems important to account for the possibility of a draw within attendance demand for football.

Another sensitivity analysis is based on the use of separate club and season fixed effects instead of club-season effects. Models (4)-(6) in Table 3.2 show the results with these separate effects. In general, not much changes compared to the results for models (4)-(6) in Table 3.1. As argued in the previous section, we think that club-season effects are preferred, since these take account of the heterogeneity of fans of different clubs with respect to intra-season differences, while separate effect do not. However, for the results, it does not seem to matter.[19]

Before we continue with a discussion on the interpretation of several results, we briefly discuss the results for seasonal uncertainty. In all specifications in Table 3.1, the size of the coefficients as well as the significance levels are much the same. Seasonal uncertainty related to the championship victory is never significant, while the uncertainty related to the qualification for the Champions League (*UEFA Champions League*) is only significant at a 10 percent level in models (3) and (6). Uncertainty related to the *UEFA Europa League* and *Relegation* is highly significant in all models. These latter results support the UOH. We further discuss the results on seasonal uncertainty, and in particular the impact of the introduction of the end-of-season play-offs, in the following section.

3.5.2 Simulations

Our baseline results showed that several variables have a statistically significant impact on the attendance rate. However, this does not tell us anything about the economic impact in terms of additional attendance. Therefore, we use the results of model (6) of Table 3.1 to get an impression of the magnitude of the effects. Since we use a linear model, the coefficients can be interpreted as marginal effects, i.e. the marginal effect measured in percentage points attendance rate. We look at the effects for a selection of variables as listed in the first column of Table 3.3. The second column gives the marginal effects. We obtain an interpretation of the effect by multiplying the coefficients with the standard deviation as given in the third column. The results are shown in column four. The ΔAR gives the percentage point change in the attendance rate that results from an additional SD. The

[19] We also performed sensitivity analyses with the use of pre-match ranks instead of the quality indicator we proposed. However, in all cases, this quality indicator outperformed the use of rankings. Furthermore, we tested with the sum and difference in quality and found some significant results, but think that the use of separate variables for the home team and the away team are more insightful.

Quality Team variable has the biggest impact. An additional SD results in 2.9 additional percentage points of attendance rate. The effect size for *Odd-Surprise Opponent* is the smallest, with an additional 0.4 percentage points of attendance rate. In general, the effects are quite small. This might have to do with the fact that attendance rates are generally quite high, with only little variation. Therefore, an alternative way to measure the effect size, is to relate the marginal effect of an additional SD to the variable part of attendance. We calculate this variable part, by taking the SD in the attendance rate for each of the 270 club-season combinations included in our analysis. Then, we take the average of these values to obtain the average standard deviation of the attendance rate (AvSDAR), which is 0.057. Column five gives the values for ΔAR/AvSDAR, i.e. the marginal effect of an additional SD, measured in terms of the average variation in the attendance rate. It follows that several effects now turn out to be rather substantial. For example, the value 0.501 for *Quality Team* means that an additional SD results in an increase of about half of the average variable part in attendance.

Table 3.3: Marginal effects and the impact on the attendance rate

Variable	Marginal effect	SD	ΔAR	ΔAR / AvSDAR
Odd-Surprise Team	0.241/100	4.80	0.012	0.203
Odd-Surprise Opponent	0.092/100	4.82	0.004	0.078
Quality Team	0.243/100	11.76	0.029	0.501
Quality Opponent	0.089/100	11.80	0.011	0.184
Match-Expectation	-0.028	0.50	0.006	0.098
Match-Expectation^2	0.012	1.63		
UEFA Europa League	0.089	0.08	0.007	0.125
Relegation	0.164	0.09	0.015	0.259

Note: SD means standard deviation; ΔAR gives the percentage point change in the attendance rate that results from an additional SD; AvSDAR is the average standard deviation of the attendance rate; the coefficients for the performance variables and quality variables are divided by 100, since these variables were divided by 100 within the models.

Next, we continue with the evaluation of additional regular league match attendance as a result of the introduction of the play-offs. Since we compare the period before and after the introduction, and given that club-specific factors might matter, captured by the club-season fixed effects, we prefer to work with a balanced panel of nine clubs that were active in the *Eredivisie* throughout the sample period. In that way, we can estimate club-specific time trends. The clubs in the balanced panel are Ajax, AZ, FC Groningen, FC Twente, FC Utrecht, Feyenoord, PSV, SC Heerenveen and Vitesse. Appendix B provides some separate details on the data for this set of clubs.

The impact of the play-offs is empirically tested by the inclusion of an interaction term for the variables *UEFA Champions League* and *UEFA Europa League*. The *UEFA Champions League* variable interacts with a dummy that takes on the value of one for the three seasons 2005/06 – 2007/08. During these seasons, play-offs for an UEFA Champions League ticket were organized. The

UEFA Europa League variable interacts with a dummy variable that takes on the value of one for all seasons after the introduction of the play-offs in 2005/06.

Panel A of Table 3.4 presents the results of this approach. Model (1) reports the results for seasonal uncertainty of our baseline model, i.e. model (6) in Table 3.1, with data from the balanced panel. It follows that, for this set of teams, *Championship* is significant at a 10 percent level, while *Relegation* is not significant. This probably has to do with the fact that the balanced panel consist of teams that perform rather well. They are generally ranked at the upper half of the league table, competing for the championship victory or European football, and not at the bottom of the table, where teams must avoid relegation. Furthermore, both the *UEFA Champions League* (at 5 percent) and *UEFA Europa League* (at 1 percent) are statistically significant and positive. We separate the effect of the play-offs in model (2) with the inclusion of the interaction terms *D*UEFA Champions League* and *D*UEFA Europa League*. For the UEFA Champions League, the coefficient is insignificant, suggesting that these play-offs did not matter for regular-league match-attendance. For the UEFA Europa League, the interaction term is highly significant, meaning that these end-of-season play-offs lead to more attendance during regular league matches. These results are robust, if we consider a subset of eight seasons in model (3), i.e. four seasons prior the introduction of the play-offs and four seasons after the introduction.

Thus far, we implicitly assumed that fans experience the qualification for the end-of-season play-offs in a similar way as a certain qualification for an European competition.[20] In the latter situation, the qualification is based on the final ranking without play-offs. In other words, the interest shifts from the qualification for UEFA competitions, to an interest in the qualification for the end-of-season play-offs. This seems reasonable, if clubs and fans experience these two types of qualification as comparable rewards for the performance during the regular league. Then, the play-offs become a sort of separate competition with its own reward, i.e. the qualification for European football. Given that many club officials and coaches nowadays state that their goal for the season is to qualify for the play-offs, the assumption, at least partly, seems to make sense. Seasonal uncertainty is modelled in line with this reasoning.

[20] It would be more precise to say, "the qualification for the end-of-season play-offs, with a possibility to qualify for a European competition", since there is always a possibility that a club earns the entry ticket to Europe after the qualification for the play-offs. However, we prefer to discuss it as the qualification for end-of-season play-off. First, because it is shorter and easier to distinguish from the other case, i.e. a direct or certain qualification for Europe. Second, and more important, because we want to stress that it has to do with the interest in the play-offs and not in particular the fact that this provides the possibility to qualify for an UEFA league.

However, if it is the qualification for European football that matters, and the end-of-season play-offs are not a separate reward at all, then seasonal uncertainty related to the introduction of the play-offs should be modelled in a differ way. As discussed by Bojke (2007) and Koning (2007), two opposing effects will emerge by the introduction of end-of-season play-offs. First, the number significant matches increases, since more final rankings provide the possibility to qualify for European football. Second, the significance per match decreases, since the number of available tickets for UEFA competitions remains unchanged. Below, we discuss how these two aspects might relate to our data.

For the full sample of 18 clubs over 15 seasons, and in the situation with play-offs, there are 2,857 matches with a non-zero significance for the Champions League variable. If we assume that there would not have been pay-offs and the runner-up always qualified for the Champions League, all else equal, the number of significant matches would have been 2,689. Similarly, the number of matches with a non-zero significance for the UEFA Europa League is 3,741. Without play-offs, and assuming that all teams up to and including rank six would have qualified up until the 2008/09 season, and after that, all teams up to and including rank number five, the number of significant matches would have been 3,578. Thus, for both variables, the introduction of the play-offs indeed increased the number of matches with a non-zero significance related to the qualification for European football. For the balanced panel of nine teams, the result is similar for the Champions League. The number of significant matches amounts 1,670 with play-offs and 1,615 without play-offs. However, for the UEFA Europa League, the number of matches is somewhat lower in the situation with play-offs (1,863), compared to the situation without play-offs (1,942). This might result from the fact that this subset of stronger teams quite easily obtains the necessary number of points to qualify for this end-of-season competition. With this exception in mind, we, in general, observe an increase of the number of significant matches in our data. As is suggested by Bojke (2007) and Koning (2007). However, the largest increase, of about 7 percent, is found for the Champions League variable with data from the full sample. This seems quite small, also compared to the results of Bojke (2007) who finds an increase of about 21 percent for the English Championship, and Koning (2007) with a predicted increase of about 50 percent, depending on the model, for the Dutch *Eredivisie*.

The opposing effect suggests that the significance per match decreases in the situation with the end-of-season play-offs. Unfortunately, Jennett's significance indicator does not contain any element that takes account of such a decrease. However, we may assume that each team in the end-of-season competition has a probability of one fourth to win. Then, we could argue that seasonal significance related to the qualification for the European competitions is only one fourth of the significance in the situation without play-offs. Obviously, the use of equal probabilities for the four teams might not be realistic, but for the sake simplicity we proceed with that assumption. We divide the variables *UEFA Champions League* and *UEFA Europa League* by four for the relevant seasons, i.e. the seasons with play-offs. Correspondingly, the interaction terms also change. We use these new variables and again test the impact of the play-offs. The results are presented in panel B of Table 3.4. Model (4) shows that none of the seasonal uncertainty variables is significant. The *p-value* for *UEFA Europa League* is 0.15, while the *p-value* for *UEFA Champions League* 0.18. In general, if we compare the results from models (1) and (4), we conclude that seasonal uncertainty seems to be less important for attendance demand, after the reduction in values for the variables that are influenced by the play-offs. However, the results for the interaction terms (models (5) and (6)) suggest that the play-offs have a positive impact on the attendance rate. Note that the coefficients for the interaction term for the UEFA Europa League, after dividing by four, are about equal to the values in panel A. The coefficient for the interaction term for the UEFA Champions League, again after dividing by four, is larger and significant in panel B compared to panel A. This seems a bit puzzling, since the values for the variable are smaller in panel B during the period with play-offs. Our general interpretation of the results in panel B is, that seasonal uncertainty does not seem to matter much, but, the play-offs had a positive effect on the attendance demand.

Table 3.4: Impact of play-offs for European football

Panel A	(1)	(2)	(3)	Panel B	(4)	(5)	(6)
Championship	0.065*	0.069	0.028	Championship	0.048	0.069	0.028
	(0.039)	(0.042)	(0.038)		(0.041)	(0.042)	(0.038)
UEFA Champions League	0.088**	0.077	0.128**	UEFA Champions League/4	0.077	0.077	0.128**
	(0.041)	(0.058)	(0.055)		(0.057)	(0.058)	(0.055)
D*Champions League		0.044	0.002	D*Champions League/4		0.408**	0.394**
		(0.074)	(0.072)			(0.187)	(0.189)
UEFA Europa League	0.097***	-0.015	-0.013	UEFA Europa League/4	0.053	-0.015	-0.013
	(0.030)	(0.020)	(0.018)		(0.037)	(0.020)	(0.018)
D*UEFA Europa League		0.169***	0.194***	D*UEFA Europa League/4		0.631***	0.736***
		(0.044)	(0.047)			(0.157)	(0.172)
Relegation	0.044	0.043	0.048	Relegation	0.027	0.043	0.048
	(0.035)	(0.036)	(0.034)		(0.038)	(0.036)	(0.034)
Observations	2,294	2,294	1,223	Observations	2,294	2,294	1,223
Censored	975	975	554	Censored	975	975	554
Club-Season FE	135	135	72	Club-Season FE	135	135	72
Opponent FE	Yes	Yes	Yes	Opponent FE	Yes	Yes	Yes

Note: Tobit regression with attendance rate as dependent variable and with the upper limit set at 0.95. Estimates are based on a balanced panel of nine teams. Robust standard errors in parentheses, clustered by club-season. Models contain all other variables included in models (6) of our baseline results (not reported). Panel B contains the results with an adjustment for the probability to qualify for European football after the introduction of the play-offs. ****p<0.01, **p<0.05, *p<0.1.*

We continue with an evaluation of the overall impact of the play-offs. From the discussion above, it follows that we could use several models, all with certain (implicit) assumptions. We suggest that it is unlikely within Dutch professional football, that clubs and fans do not value the qualification for end-of-season play-offs. Also, we do not think that all clubs and fans will value the qualification for end-of-season play-offs always and exactly the same as a certain qualification for European football (i.e. the situation without play-offs). Thus, a better understanding on this aspect is needed and something to deal with in future research. For now, we suggest that the former assumption fits better with the current situation in Dutch football, i.e. the qualification for the play-offs is valued by fans. Therefore, we proceed with the results of model (2) of Table 3.4. These results are used to obtain model predictions. For each observation, we obtain a fitted value, taking into account the upper limit of 0.95. Then, we also calculate a value without taking account of the interaction terms, as if there would have never been play-offs. The results are shown in Figure 3.2, where we plot the mean predicted values by season for both situations with and without play-offs.

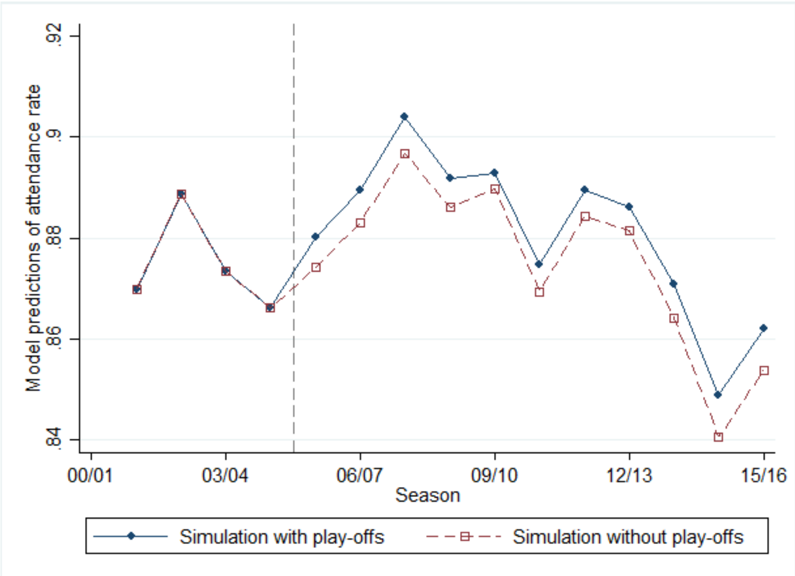


Figure 3.2: Model predictions of mean attendance rate with and without play-offs. The reference line indicates the introduction of play-offs in the season 2005/06.

Obviously, for the first four seasons the graphs coincide. From the seasons 2005/06 onwards, after the introduction of the play-offs, we observe very small differences. This suggests that the play-offs have made a rather stable contribution to the within-season variation of the stadium attendance rate during regular league matches. It also suggests that between-season variation, and in particular the increase in the average attendance over time (as observed in Figure 3.1), cannot be attributed to the play-offs.

Of course, the end-of-season play-off matches themselves also attract attendance. This additional attendance should be considered in an evaluation of the effectiveness of the play-offs on stadium attendance. In Figure 3.3, we plot the attendance rate during the play-off matches against the attendance rate during the same match in the regular league. The special cases for FC Groningen and FC Twente are indicated, because these play-off matches were played in a different stadium than the regular league match. Dots above the 45° line indicate that attendance was higher during the regular league match. Although we observe a lot of dots rather close to the line, the majority lies within the upper left part of the figure. While match uncertainty, the quality of the competing teams as well as seasonal uncertainty should be relatively high for the play-off matches, fans seem to be less likely to attend them, compared to regular league matches. This might have to do with the fact that season tickets holders, in general, have to pay extra for these play-offs.

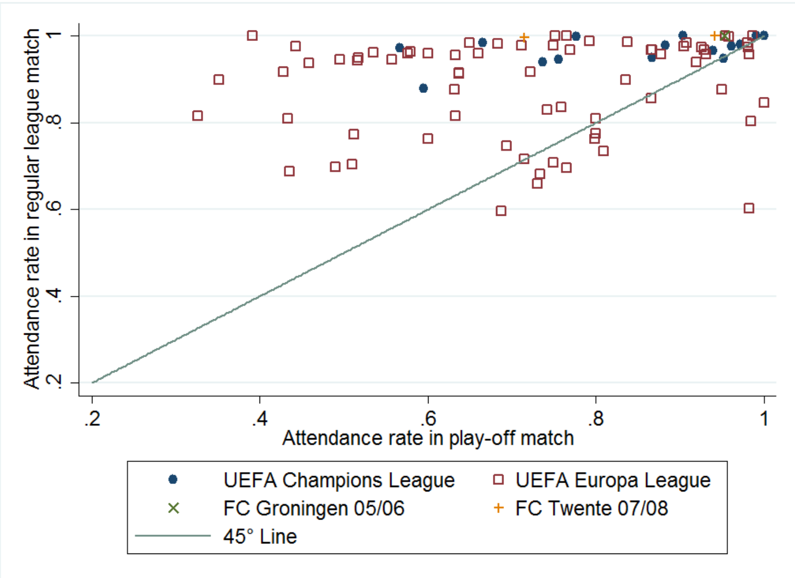


Figure 3.3: Scatterplot of attendance during the play-offs and regular league matches.

Besides this modest interest in the play-off matches by fans, there are some other drawbacks of the system. The play-off matches are generally seen as an unwanted extension of the season by clubs, managers and players. This holds especially in the seasons when there is a UEFA European Championship or FIFA World Cup coming up, for which preparations start immediately after the season. Furthermore, it can be argued that it is not always the strongest team that wins the play-offs and, thus earns a ticket for European football in next season. Koning (2007) found that this probably is the case. For Dutch football in general, it would be better if the best teams represent the Netherlands in European competitions. They have the best perspective to survive the group stage and maybe a few rounds in these competitions, earning points for the UEFA rankings for club competitions.

Furthermore, such international matches are valuable for the development of young and talented players who might be selected for national teams. In general, these players play for the better teams.

3.5.3 Changes in attendance over time

The focus so far has been on within-season variation in attendance demand. Between-season – between-club variation is picked by club-season fixed effects. The question we address in this subsection is to what extent the between-season variation of attendance for a particular club is influenced by stadium capacity. Figure 3.4 shows that, on a club level, there is a strong correlation between stadium capacity and stadium attendance.

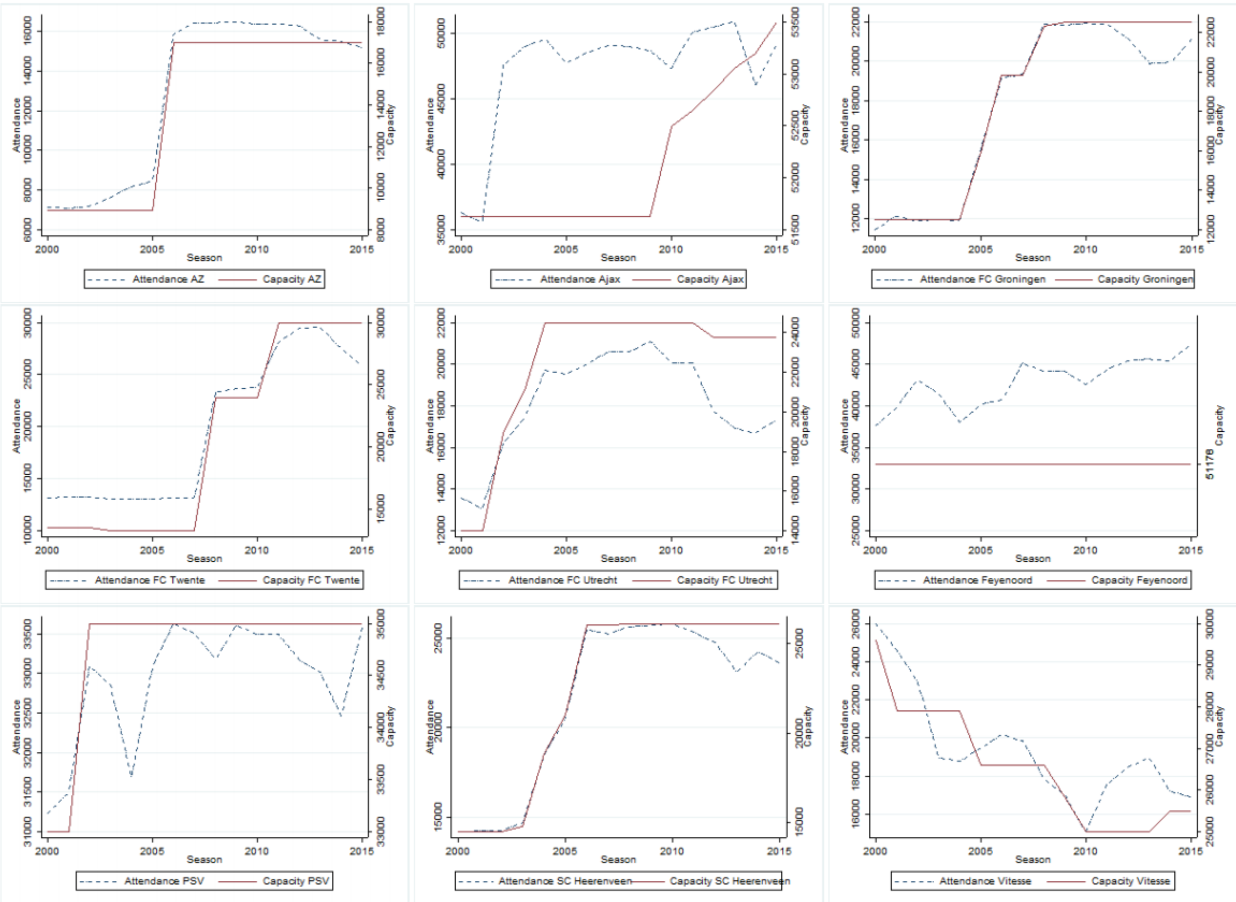


Figure 3.4: Stadium capacity and attendance, nine clubs; 2000/01-2015/16

This correlation is especially high in situations where the expansion of capacity was relatively large. AZ expanded their stadium capacity from 9,000 up to 17,000 in 2006, while FC Groningen enlarged their stadium capacity, in two major steps, from 12,500 up to 22,000. The first expansion took place during the season 2005/06, when they moved to a new stadium. The stadium of FC Twente was also scaled-up substantially in two steps, from 13,000 via 24,000 up to 30,000. FC Utrecht expanded their stadium capacity in a few steps from 14,000 up to 24,000, while SC Heerenveen enlarged their stadium capacity from 14,500 up to 26,000. Relatively minor expansions took place in the stadiums of Ajax and PSV. The capacity of the Feyenoord stadium remained unchanged during the sample period, while the stadium capacity of Vitesse decreased from 30,000 down to 25,000.

It follows that a strong expansion of stadium capacity almost immediately led to a substantial increase in attendance. Furthermore, in their study on novelty effects of German football stadiums, Feddersen, Maenning and Borcherding (2006) discuss the possibility for capacity effects, i.e. expansions because of excess demand. They suggest that a capacity effect emerges if the average attendance rate is more than 90% in the three seasons prior to the decision to renovate and upgrade the current stadium, or to build a new one. Within our balanced panel, we can identify eleven substantial expansions in stadium capacity, i.e. expansions with more than 500 seats, for which the average attendance rate for the last three seasons is available. In only one case, the value is below the threshold level of 90% (i.e. FC Utrecht in the season 2004/05, with a value of 87%) while ten cases would be defined as capacity effect. Thus, in general, most expansions seem to have been triggered by excess demand for (season) tickets.

An alternative explanation is that the new, enlarged and renovated stadiums attracted new attendants that were not interested to go to the stadium before the improvement of facilities. Thus, attendance increases, because of the new or renovated stadium. This phenomenon is known as the novelty effect (or honeymoon effect) and has been studied numerous times. For example, Coates and Humphreys (2005) study the novelty effect for three major US sports leagues. They find quite substantial effects, with the largest effect for the MLB, then for the NBA and finally, the NFL. They also find that durations of the effect differ for the different leagues. Within the NFL, the novelty effects only lasts for about five years, while respectively eight and nine years for the other leagues. Feddersen, Maenning and Borcherding (2006) study the novelty effect for German football stadiums. They find a rather small effect (compared to US sports) within the first five seasons after the opening of a new (or renovated) stadium. Finally, Love *et al.* (2013) study the novelty effects in US Major League Soccer (MLS). In particular, they look for shifts in average attendance for teams that move from multipurpose venues to soccer-specific stadiums. Although they do not control for other elements that may affect attendance, they find substantial novelty effects for at least the first three seasons. Generally, given the rather robust results from previous studies, it seems likely that our sample also exhibits at least some novelty effects. However, the jumps in attendance following an expansion of capacity suggest that the explanation of excess demand is more likely. Still, both explanations seem reasonable and might be complementary. Therefore, we should be cautious with causal conclusions.

Table 3.5: Parameter estimates club-season variation in attendances

	(1)	(2)	(3)
Log Stadium capacity	0.987***	0.971***	0.961***
	(0.026)	(0.047)	(0.043)
Ranking end-of-season			-0.062***
			(0.023)
Club FE (9)	Yes	Yes	Yes
Season FE (16)	No	Yes	Yes

Note: Dependent variable is logarithm of attendance per club-season. All estimates contain 144 observations. Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.*

To quantify the relationship between capacity and attendance, we estimate a simple model with the logarithm of attendance per club-season as the dependent variable and the logarithm of capacity as one of the explanatory variables. The relevant parameter estimates are presented in Table 3.5.
The first column shows, that if we only include club fixed effects, then the effect of *Log Stadium capacity* is 0.99, which is not significantly different from unity. A one percent increase in stadium capacity also increases attendance by one percent. The second column shows that this parameter estimate hardly changes, if we also include season fixed effects. In the third column, we add the ranking at the end of the season. This has a significantly negative effect on attendance, but the relationship between stadium capacity and attendance is hardly affected.[21] Again, these very high correlations suggest that attendance directly increases, in about equal portions, after the enlargement of the stadium. As discussed, this may follow from the supply of new seats and facilities, i.e. the novelty effect. Alternatively, this could result from the excess demand for (season) tickets. It is unclear to what extent supply and demand contribute to the result. We suggest that the latter explanation seems plausible. However, since our analysis lacks causality, this remains a question open for future research.

3.6 Discussion and conclusion

We investigate the relation between outcome uncertainty and within-season variation in stadium attendance for sixteen seasons of the highest tier of Dutch professional football. Using different measures of match uncertainty, we find evidence that contradicts with the UOH. Instead, results suggest that fans have reference-dependent preferences with loss aversion. Although we cannot rule

[21] Note that a higher value of the ranking at the end of the season means that a team did worse compared to a lower value.

out any preference for uncertain outcomes, the results suggest that loss aversion dominates. However, no home win preference is found. Furthermore, the influence of match uncertainty disappears after controlling for opponent fixed effects in our baseline model. In general, we conclude that team characteristics are more important for the determination of stadium attendance, than behavioural economic explanations regarding the outcome of the match (i.e. a preference for outcome uncertainty, loss aversion and a home win preference). In particular, the in-season performance as well as the quality of both the home team and the away team have a positive effect on the attendance rate. We introduced a new indicator to proxy for the quality of teams, based on bookmaker odds, that works quite well within the attendance demand model. A sensitivity analysis shows that our results are robust for specifications with separate club and season fixed effects instead of our preferred specification with club-season effects. Furthermore, this analysis reveals that our variable for match-expectations outperforms the home win probability as indicator of match uncertainty. This stresses the importance to account for the probability for a draw in attendance demand models within football. As to seasonal uncertainty, we find significant results for uncertainty related to the qualification for the UEFA Europa League and seasonal significance related to relegation. Both are in line with the UOH. The insignificant findings for the other seasonal uncertainty variables might result from the fact that, in general, only a few teams perform well enough to remain a serious contestant for the title or to qualify for the UEFA Champions League until the very final stages of the season. Both of these variables, as well as the variable for UEFA Europa League, are significant for a balanced panel of nine teams that were active at the highest level in all seasons considered. This result is in line with the fact that these teams generally perform rather well and, thus, experience seasonal uncertainty with regard to the top of table rankings and the corresponding options for qualification.

Our preferred model is used to obtain an impression of the magnitude of several effects. In general, the marginal effects are small. However, in relation to the variable part of the attendance rate, being small itself, we find some rather substantial results. Furthermore, we investigate the impact on regular league match-attendance of the introduction of the play-offs in the season 2005/06. We find that the UEFA Europa league play-offs have a positive effect on stadium attendance during regular league matches. Nevertheless, the magnitude of the effect is small. Furthermore, our final part of analysis reveals that the play-offs did not drive the increase in average attendance figures over seasons. Between-season variation in attendance is correlated with stadium capacity. The jumps in attendance after an increase of the capacity suggest that these expansions were necessary to meet excess demand. Given some other drawbacks of the play-offs, it seems reasonable to question their overall effectiveness.

In general, we find that stadium attendance rates are high within the Dutch *Eredivisie*. Thus, suggesting that the current policy of clubs and the sports body does not need to be changed. Furthermore, given the small marginal effects, one might question the effectiveness of any policy measure related to the aspects investigated in this study. There seems to be a vast majority of fans that simply attends any match, irrespective of the conditions, competing teams and the season. As long as this majority keeps returning to the stadium, there does not seem to be any problem. However, both figures 3.1 and 3.2 show some reduction in the average attendance rate for the recent seasons. If this continues over the following seasons, clubs may feel the need to do something. Then, it is useful to understand why people (do not) attend the stadium, so that one can formulate clear policy advice. Our results may serve as a basis. Future research could explain why people might be interested to attend matches against certain opponents, i.e. the opponent effect.

Appendix B: Information about our data

Our data on match date, competing teams, match results, stadium capacities and attendance figures are collected from various internet sources, sports magazines and newspaper archives. The attendance figures are typically a reported number, from which it is not possible to separate season-ticket holders and pay-at-the-gate attendance. Furthermore, the reported data might be rounded to e.g. 100, which is likely to hold true as we observe several spikes of frequencies at those numbers. Data on the general weather conditions are obtained from the KNMI, the Royal Dutch Meteorological Institute. The betting odds are from bookmaker agencies William Hill (97 percent), Ladbrokes (2 percent) and others (1 percent). Still, no data was available for two matches. From the data collected, we create the set of variables listed in Table B1.

In Table B2, some descriptive statistics are presented. Since we need previous seasons bookmaker odds for the construction of the quality variables, and we lack odds prior to the season 2000/01, this season is left out of the analyses. Thus, we are left with 15 seasons with 18 clubs and 17 home matches per club, resulting in 4,590 observations. For 186 observations, we found a reported number of attendance that was (slightly) higher than our documented stadium capacity. For these cases, we assumed attendance to be the same as capacity. During two matches, no fans were allowed to attend as a punishment by the Dutch football association. These two matches are left out of the analyses. Furthermore, bookmaker odds are missing for two matches, which results in 4,586 observations for the match uncertainty measures that are based on these bookmaker odds. A total of 1,980 observations is censored at the upper limit of 95 percent of capacity (41 percent of observations is censored). The right hand panel of Table B2 contains descriptive statistics for a subsample of 9 teams that were active in the *Eredivisie* throughout the sample period. This balanced panel contains 2,294 observations, with 975 censored observations at 95 percent of capacity (43 percent of the observations). We find a rather high standard deviation and difference between the minimum and maximum value for attendance, which can easily be explained by the fact that all clubs are grouped together for these values. After controlling for capacity in the attendance rate variable, the deviations are less severe.

Furthermore, it is noteworthy that the Odd-Surprise variables are slightly positive on average and fluctuate in both directions. The quality variables cannot become negative, by definition, and the mean values are approximately 47. This is about the number of points that teams typically obtain if they finish in the middle of the league, thus representing average quality. We also find that the mean values for the Theil-index and PPG are rather high (low) in comparison to their minimum and maximum values, which is less the case for the Home win probability and Match-expectation. Finally, the mean values for seasonal uncertainty are rather low, which can be explained by the fact that seasonal uncertainty develops exponentially. For the first match the value is equal to 0.03 in all cases. For a given match later in the season, the value is equal to 0.5 when only two matches should be won, while it only reaches unity in case only one match needs to be won to obtain some end-of-season achievement.

In Figure B1 we plot the relevant number of points necessary to obtain an end-of-season achievement, i.e. pt_k in equation (3). The vertical reference-line marks the introduction of the end-of-season play-off scheme. The line for relegation remains flat for the entire sample period. For the UEFA Europa League, a stable number of points is necessary to qualify until the introduction of the play-offs. After the introduction, again a stable number of points is necessary for qualification, now for the end-of-season play-offs. The line representing 'championship' exhibits a higher degree of variability, although it remains somewhere around 80 points. No stable pattern is found for the UEFA Champions League. This can be explained by changes in the number of direct tickets for this tournament, as well as the introduction of the play-offs during the sample period. Only one extra team, besides the champion, was allowed to qualify from the season 2001/02 onwards, as a result of the rather poor performances of the Dutch teams in Europe in the preceding seasons. However, the introduction of the play-off scheme in the 2005/06 season made it possible for multiple teams to obtain this qualification in an end-of-season competition. Teams ranked second up to fifth were competing for this second entry ticket during the 2005/06 – 2007/08 seasons. The playoffs were abandoned after the 2007/08 season, after which the runner up in the league directly qualified for the Champions League.[1] Again, we observe a rather flat line from that moment onwards, with two exceptions, in which both the champion as well as the runner-up obtained more than 80 points, which is rather high. In general, taking the changes in rules and regulations regarding qualifications into account, we find that the necessary number of points, i.e. pt_k, is rather stable over seasons. Especially for relegation and the UEFA Europa League (before and after the introduction of play-offs), which suggests that the assumption of a known final table is comparable to expectations being based on previous years' results.

[1] Only once out of three times, the runner-up won the play-offs for the lucrative Champions League (Ajax in the 2006/07 season). In the other two seasons, the club ranked fourth in the regular league won the ticket. Many commentators considered this as a result that would not enhance the position of Dutch clubs in competitions organized by the UEFA. Furthermore, note that, when the play-offs for the UEFA Champions League were canceled, the runner-up in the league qualifies for a start-of-season preliminary round of the UEFA Champions League, but for convenience, we do not make such a distinction here.

Table B3 provides pairwise correlations. Without going into detail, we find many variables that are significantly correlated to the attendance variables. Those measuring uncertainty of outcome are of main interest. The match variables all significant, but the sign of the coefficients is different from the *a priori* hypotheses. For seasonal uncertainty, we find positive correlations for top-of-the-table rankings, i.e. Championship and UEFA Champions League, while negative correlations for uncertainty related to UEFA Europa League and Relegation. In general, these results do not support the UOH.

Table B1: Description of variables

Variable	Description
Attendance (dependent variable)	
Attendance Rate	Occupancy rate of the stadium, calculated as the ratio of attendance divided by stadium capacity
Log Attendance	Logarithm of attendance
Opponent	
Derby	Dummy with value 1 if teams are from the same province (+)
Match day	
Weekday	Dummy with value 1 if match was played on a Monday, Tuesday, Wednesday or Thursday. Value is zero in case of a national holiday (-)
Weather conditions:	
Temperature	Daily mean temperature measured in 0.1 degrees Celsius in De Bilt on match-day (+)
Precipitation	Daily precipitation, amount in 0.1 mm in De Bilt on match-day (-)
Performance	
Odd-Surprise Team	Cumulative-Surprise based on bookmaker odds, calculated as the in-season sum of surprises. A single match surprise is calculated as the actual number of points obtained minus the expected number of points, where the expected number of points is based on the probabilities set by bookmakers for a home win, a draw and an away win (+)
Odd-Surprise Opponent	Cumulative-Surprise of the opponent, calculated in the same way as cumulative surprise for the Team (+)
Quality	
Quality Team	Expected quality of the team, measured by the sum of the expected number of points in the previous 34 matches. Expectations are based on bookmaker odds. In case a team was promoted, and, thus, no information about previous season is available, we take the expected number of points of the team that relegated and was replaced by the promoted team (+)
Quality Opponent	Expected quality of the opponent, calculated in the same way as the expected quality for the Team (+)
Ranking Team	Pre-match rank of the team (-)
Ranking Opponent	Pre-match rank of the opponent (-)
Match-UO	
Theil	Theil-index, calculated as: $\sum_{i=1}^{3} p_i \ln(\frac{1}{p_i})$ with p_i being either the home win probability, the probability for a draw or the away win probability. The index is increasing when probabilities become more equal, with a value close to zero in case of a high probability for one of the outcomes and ln(3) in case of equal probabilities (+)
PPG	Points-per-game measure, calculated as: $PPG_{ijk} = \lvert HA_k + PPG_{ik} - PPG_{jk} \rvert$ in which, i denotes the home team, j indicates the away team and k refers to the season. The points-per-game match uncertainty measure PPG_{ijk} is the absolute value of home advantage (HA_k) plus the number of points per game of the home team (PPG_{ik}) minus the number of point per game of the away team (PPG_{jk}). Home advantage is measured as the difference between the points per game won by all home teams and all away teams in previous season. The point per game values of the home and away team are measured as the number of points per game obtained in the current season. With this measure, match uncertainty increases with decreasing values (-)
Home win probability	Home win probability based on bookmaker odds (-)
Home win probability^2	Home win probability squared (+)
Match-Expectation	Expected number of points for the home team, calculated as: $p_{win} * 3 + p_{draw} * 1 + p_{lose} * 0$, where probabilities are based on bookmaker odds. (-)
Match-Expectation^2	Match-Expectation squared (+)
Seasonal-UO Either being "Championship Significance", "UEFA Champions League Significance", "UEFA Europa League Significance" or " Relegation Significance"	Significance is defined as: $s_{ijk} = \frac{1}{m_{jk} - n_{ijk}}\ if\ pp_{ijk} \geq pt_k > pc_{ijk};\ \ s_{ijk} = 0\ if\ pp_{ijk} < pt_k;\ \ s_{ijk} = 0\ if\ pc_{ijk} \geq pt_k$ in which i denotes the match, j indicates the club and k refers to the season. Then, significance s_{ijk} is given by the reciprocal of the total number of matches in the season (m_{jk}) minus the number of matches already played (n_{ijk}) prior to match i. This is only the case if the potential number of points for a specific team (pp_{ijk}) is larger or equal to the total number of points needed to obtain the predefined total number of points. This total is constant throughout the season and represented by pt_k, which should be larger than the current number of points (pc_{ijk}) prior to a match. In case the potential number of points is smaller than the total number of points, significance becomes zero (+)

Note: National holidays include Easter, Queen's Birth Day / King's Birth Day, Ascension Day, Pentecost and Christmas. In the Netherlands, there is national holiday on Sunday and Monday during Eater and on Sunday and Monday during Pentecost. Furthermore, there is a national holiday on the 26th of December, the second day of Christmas. De Bilt is a municipality centrally located in the Netherlands where the KNMI (Royal Dutch Meteorological Institute) is located. Precipitation was measured as being -1 for values <0.05, but was set to zero in our sample. Both temperature and precipitation measures were divided by 10 to make interpretation easier. The sign (+ or -) in parentheses indicates the expected sign of the coefficient.

Table B2: Descriptive statistics

	All teams					Balanced panel of 9 teams				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Attendance	4,588	18,191.12	12,086.46	2,258	53,052	2,295	26,128.41	12,306.86	6,057	53,052
Attendance Rate	4,588	0.88	0.12	0.30	1	2,295	0.89	0.11	0.42	1
Log Attendance	4,588	9.61	0.65	7.72	10.88	2,295	10.06	0.48	8.71	10.88
Derby	4,588	0.09	0.29	0	1	2,295	0.06	0.24	0	1
Weekday	4,588	0.07	0.25	0	1	2,295	0.06	0.24	0	1
Temperature	4,588	0.89	0.56	-1.21	2.64	2,295	0.89	0.56	-1.21	2.64
Precipitation	4,588	0.23	0.42	0	4.00	2,295	0.23	0.43	0	4.00
Odd-Surprise Team	4,588	0.06	4.80	-17.36	19.98	2,295	1.21	4.96	-17.36	19.98
Odd-Surprise Opponent	4,588	0.09	4.82	-18.31	19.29	2,295	0.10	4.80	-17.09	18.93
Quality Team	4,588	46.96	11.76	23.52	76.95	2,295	55.25	10.40	35.71	76.95
Quality Opponent	4,588	46.71	11.80	23.17	76.38	2,295	46.14	11.67	23.17	76.28
Ranking Team	4,588	9.37	5.30	1	18	2,295	6.60	4.51	1	18
Ranking Opponent	4,588	9.13	5.32	1	18	2,295	9.21	5.29	1	18
Theil	4,586	0.98	0.13	0.44	1.1	2,294	0.93	0.15	0.44	1.1
PPG	4,588	0.83	0.63	0	3.94	2,295	1.01	0.68	0	3.94
Home win probability	4,586	0.46	0.18	0.06	0.88	2,294	0.55	0.17	0.11	0.88
Home win probability^2	4,586	0.24	0.17	0	0.78	2,294	0.33	0.18	0.01	0.78
Match-Expectation	4,586	1.63	0.50	0.31	2.73	2,294	1.87	0.46	0.53	2.73
Match-Expectation^2	4,586	2.91	1.63	0.10	7.44	2,294	3.73	1.65	0.28	7.44
Championship	4,588	0.03	0.05	0	1	2,295	0.04	0.07	0	1
Champions League	4,588	0.04	0.07	0	1	2,295	0.05	0.09	0	1
UEFA Europa League	4,588	0.06	0.08	0	1	2,295	0.06	0.09	0	1
Relegation	4,588	0.05	0.09	0	1	2,295	0.03	0.05	0	1

Note: the balanced panel consist of nine clubs that are active in the *Eredivisie* in all seasons considered. Since bookmaker odds are missing for two matches (one for the balanced panel), *Theil, Home win probability* and *Match-Expectation* contain two (one) less observation than the other variables.

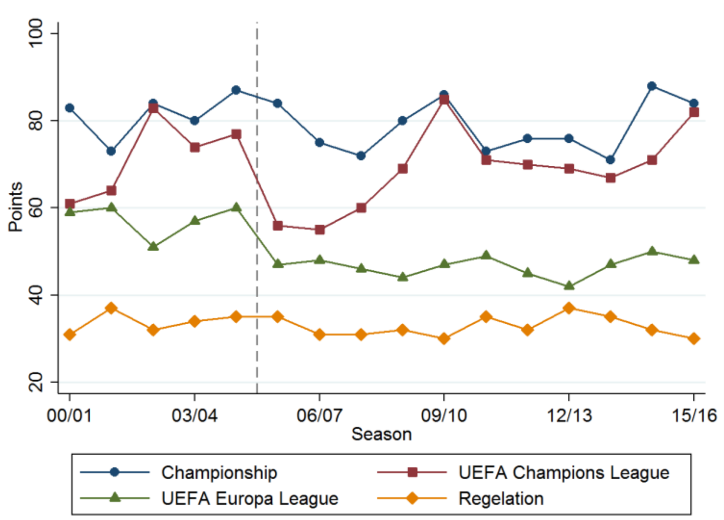


Figure B1: Points needed to obtain some end-of-season achievement. The reference line indicates the introduction of play-offs in the season 2005/06.

Table B3: Pairwise correlations

	All Teams	1	2	3	4	5	6	7	8	9	10	11	12
1	Attendance	1											
2	Attendance Rate	0.1429*	1										
3	Log Attendance	0.9329*	0.1418*	1									
4	Derby	-0.0234	0.0689*	-0.0412*	1								
5	Weekday	-0.0097	-0.0544*	-0.0129	-0.0207	1							
6	Temperature	-0.0115	-0.0303*	-0.0145	-0.0039	-0.0769*	1						
7	Precipitation	-0.0029	-0.0176	0.0013	0.0193	-0.0115	0.0502*	1					
8	Odd-Surprise Team	0.2153*	0.1121*	0.2467*	-0.0295*	0.0010	-0.0104	0.0043	1				
9	Odd-Surprise Opponent	0.0040	0.0664*	0.0030	-0.0131	0.0007	0.0106	-0.0015	-0.0341*	1			
10	Quality Team	0.8163*	0.1394*	0.7938*	-0.0206	-0.0082	-0.0011	0.0050	0.2940*	-0.0139	1		
11	Quality Opponent	-0.0212	0.1779*	-0.0136	-0.0152	0.0045	0.0114	-0.0021	-0.0361*	0.2894*	-0.0813*	1	
12	Ranking Team	-0.5774*	-0.1388*	-0.5816*	0.0155	0.0059	-0.0663*	-0.0086	-0.6686*	0.0102	-0.7061*	0.0208	1
13	Ranking Opponent	-0.0082	-0.1249*	-0.0113	0.0197	0.0044	-0.1031*	-0.0035	0.0310*	-0.6724*	0.0363*	-0.7177*	0.0615*
14	Theil	-0.5178*	-0.1262*	-0.4660*	-0.0455*	-0.0055	0.0151	-0.0136	-0.2309*	0.1819*	-0.5901*	0.2364*	0.4441*
15	PPG	0.3403*	0.0194	0.3140*	0.0254	-0.0094	0.0844*	0.0130	0.3800*	-0.3370*	0.4332*	-0.3173*	-0.4671*
16	Home win probability	0.5371*	-0.0138	0.5225*	-0.0069	-0.0085	-0.0157	0.0074	0.3330*	-0.3246*	0.6840*	-0.7020*	-0.5419*
17	Home win probability^2	0.5778*	0.0236	0.5513*	0.0072	-0.0058	-0.0174	0.0092	0.3335*	-0.3116*	0.7171*	-0.6322*	-0.5616*
18	Match-Expectation	0.5230*	-0.0234	0.5116*	-0.0101	-0.0090	-0.0154	0.0074	0.3317*	-0.3258*	0.6714*	-0.7154*	-0.5341*
19	Match-Expectation^2	0.5636*	0.0081	0.5424*	0.0016	-0.0066	-0.0167	0.0095	0.3361*	-0.3192*	0.7070*	-0.6646*	-0.5572*
20	Championship	0.2906*	0.0712*	0.2594*	-0.0253	-0.0016	0.0966*	-0.0064	0.2928*	-0.0257	0.3549*	-0.017	-0.3309*
21	Champions League	0.3031*	0.0729*	0.2725*	-0.0285	0.0161	0.002	-0.0029	0.2734*	-0.0068	0.3080*	-0.0295*	-0.2998*
22	UEFA Europa League	-0.0298*	0.023	0.0152	-0.0138	0.0129	-0.0838*	0.0024	0.0701*	0.0266	-0.0286	0.0196	-0.0274
23	Relegation	-0.2048*	0.0153	-0.2039*	0.0202	-0.0141	-0.0328*	-0.0216	-0.2705*	0.0059	-0.2580*	0.0085	0.3133*

		13	14	15	16	17	18	19	20	21	22	23
13	Ranking Opponent	1										
14	Theil	-0.2522*	1									
15	PPG	0.4418*	-0.6225*	1								
16	Home win probability	0.5770*	-0.6175*	0.5740*	1							
17	Home win probability^2	0.5342*	-0.7678*	0.6311*	0.9774*	1						
18	Match-Expectation	0.5842*	-0.5711*	0.5555*	0.9981*	0.9635*	1					
19	Match-Expectation^2	0.5557*	-0.7102*	0.6138*	0.9920*	0.9957*	0.9834*	1				
20	Championship	0.0055	-0.2665*	0.2431*	0.2539*	0.2802*	0.2455*	0.2694*	1			
21	Champions League	0.0239	-0.2215*	0.1829*	0.2321*	0.2500*	0.2262*	0.2431*	0.2852*	1		
22	UEFA Europa League	-0.0002	0.0587*	-0.0774*	-0.0168	-0.0307*	-0.0134	-0.0245	-0.1420*	-0.0950*	1	
23	Relegation	0.0161	0.1510*	-0.1836*	-0.1791*	-0.1860*	-0.1752*	-0.1837*	-0.1600*	-0.1799*	-0.0851*	1

	Balanced panel of 9 teams	1	2	3	4	5	6	7	8	9	10	11	12
1	Attendance	1											
2	Attendance Rate	0.1580*	1										
3	Log Attendance	0.9694*	0.1528*	1									
4	Derby	0.1016*	0.0129	0.1044*	1								
5	Weekday	0.0039	-0.0259	-0.0016	0.0047	1							
6	Temperature	-0.0177	-0.0525*	-0.0254	-0.0392	-0.0688*	1						
7	Precipitation	-0.0154	-0.0154	-0.0118	-0.0078	-0.0255	0.0632*	1					
8	Odd-Surprise Team	0.0524*	0.0959*	0.0399	-0.0045	0.0159	-0.0785*	-0.0004	1				
9	Odd-Surprise Opponent	0.0037	0.0692*	0.0025	-0.0178	-0.0169	0.0173	-0.0295	-0.0285	1			
10	Quality Team	0.7138*	0.2734*	0.6963*	0.1395*	0.0111	0.0013	0.009	0.1355*	-0.0261	1		
11	Quality Opponent	-0.0021	0.1477*	0.0016	-0.0756*	-0.0098	-0.0097	-0.0066	-0.0267	0.2923*	-0.0714*	1	
12	Ranking Team	-0.4604*	-0.2242*	-0.4439*	-0.0821*	-0.0342	0.0286	0.0000	-0.6034*	0.0233	-0.6378*	0.0081	1
13	Ranking Opponent	-0.021	-0.1133*	-0.0222	0.0569*	0.0257	-0.0948*	0.0031	0.0462*	-0.6831*	0.0357	-0.7109*	0.0299
14	Theil	-0.4416*	-0.1064*	-0.4310*	-0.1443*	-0.0228	0.0205	-0.0083	-0.2321*	0.3128*	-0.6344*	0.5425*	0.4892*
15	PPG	0.2646*	0.0390	0.2491*	0.0875*	0.021	0.0315	0.0081	0.3912*	-0.4521*	0.4109*	-0.4353*	-0.4745*
16	Home win probability	0.4121*	0.0552*	0.3963*	0.1236*	0.0119	-0.0123	0.0156	0.2297*	-0.3430*	0.6242*	-0.7375*	-0.4766*
17	Home win probability^2	0.4366*	0.0752*	0.4217*	0.1350*	0.0149	-0.0159	0.0133	0.2380*	-0.3442*	0.6498*	-0.6971*	-0.4974*
18	Match-Expectation	0.4034*	0.0502*	0.3867*	0.1198*	0.0125	-0.0112	0.0168	0.2271*	-0.3412*	0.6142*	-0.7463*	-0.4687*
19	Match-Expectation^2	0.4271*	0.0665*	0.4113*	0.1301*	0.0157	-0.014	0.0153	0.2364*	-0.3451*	0.6403*	-0.7186*	-0.4902*
20	Championship	0.2120*	0.1122*	0.1990*	-0.0049	0.0127	0.0388	-0.0183	0.2809*	-0.0382	0.3273*	-0.0126	-0.3032*
21	Champions League	0.2238*	0.1128*	0.2093*	-0.0036	0.0363	-0.0195	-0.0172	0.2387*	-0.0166	0.2320*	-0.0325	-0.2547*
22	UEFA Europa League	-0.1041*	-0.0092	-0.0860*	-0.0182	-0.0021	-0.0695*	-0.0025	-0.0628*	0.0218	-0.1638*	0.0392	0.0978*
23	Relegation	-0.1888*	-0.1063*	-0.1937*	-0.0162	-0.0187	-0.0098	0.0116	-0.3193*	0.0369	-0.2671*	0.0138	0.3362*

		13	14	15	16	17	18	19	20	21	22	23
13	Ranking Opponent	1										
14	Theil	-0.4925*	1									
15	PPG	0.5913*	-0.6766*	1								
16	Home win probability	0.6041*	-0.8556*	0.6483*	1							
17	Home win probability^2	0.5866*	-0.9310*	0.6788*	0.9853*	1						
18	Match-Expectation	0.6070*	-0.8284*	0.6382*	0.9985*	0.9751*	1					
19	Match-Expectation^2	0.5976*	-0.9004*	0.6698*	0.9953*	0.9966*	0.9895*	1				
20	Championship	0.0194	-0.2413*	0.2097*	0.2176*	0.2336*	0.2122*	0.2265*	1			
21	Champions League	0.0401	-0.1805*	0.1442*	0.1730*	0.1825*	0.1702*	0.1786*	0.2312*	1		
22	UEFA Europa League	-0.0160	0.1307*	-0.1404*	-0.1190*	-0.1271*	-0.1164*	-0.1240*	-0.1627*	-0.1236*	1	
23	Relegation	-0.0068	0.1747*	-0.1807*	-0.1898*	-0.1911*	-0.1861*	-0.1890*	-0.1447*	-0.1799*	-0.0416*	1

Note: * indicates significance at 0.05

Appendix C: Overview of previous studies

This appendix provides a summary overview of studies presented and discussed in section 3.2. Table C1 gives a summary of studies that focused on match uncertainty. Table C2 focuses on studies on seasonal uncertainty.

Table C1: Overview of studies with match uncertainty

Match uncertainty	Country-Division	Seasons	Dependent Variable	Match uncertainty	UOH	Comment
Hart, Hutton and Sharot (1975)	England-1	1969/70 – 1971/72	Log of match-day stadium attendance	1) Log of home team rank 2) Log of away team rank 3) Log of difference in rank	1) No* 2) No* 3) No	Separate analyses for a selection of four teams
Crains (1987)	Scotland-1	1971/72 – 1979/80	Match-day stadium attendance	1) Home team rank 2) Away team rank	1) Yes* 2) Yes*	Separate analyses for a selection of three teams; also includes seasonal uncertainty
Janssens and Késenne (1987)	Belgium-1	1982/83	Match-day stadium attendance	1) Average goals home team 2) Average goals away team 3) Average points home team 4) Average points away team	1) No* 2) No* 3) No* 4) No*	Also includes seasonal uncertainty; these measures are included here as match uncertainty, but the authors do not discuss them as such
Peel and Thomas (1988)	England-1/4	1981/82	Log of match-day stadium attendance	1) Home team rank 2) Away team rank 3) Home team win probability	1) No* 2) No* 3) No	Separate analyses for the different divisions
Dobson and Goddard (1992)	England-1/4	1989/90 – 1990/91	Log of match-day stadium attendance, distinguishing standing and seating	1) Log of home team rank 2) Log of away team rank	1) No* 2) No*	Analyses based on a set of clubs that (were able to) provide survey information; also includes seasonal uncertainty
Peel and Thomas (1992)	England-1/4	1986/87	Log of match-day stadium attendance	1) Home team rank 2) Away team rank 3) Home team win probability, also squared	1) No* 2) No* 3) Partial	Separate analyses for the different divisions; formulate the Theil-index, but use home win probability
Wilson and Sim (1995)	Malysia-1/2	1989/90 – 1991/92	Log of match-day stadium attendance	1) Absolute point difference between teams, also squared	1) No	Non-paying spectators and season ticket holders are excluded from the analyses; also includes seasonal uncertainty
Baimbridge, Cameron and Dawson (1996)	England-1	1993/94	Log of match-day stadium attendance	1) Difference in league rank, also squared	1) No	Also includes seasonal uncertainty
Peel and Thomas (1996)	Scotland-1/3	1991/92	Difference in stadium attendance figures in repeat fixtures	1) Difference in home team rank 2) Difference in level of home win probability, also squared 3) Difference in Theil-index	1) No 2) Partial 3) No	The Scottish league contains repeat fixtures between the same home and away teams during a season
Baimbridge (1997)	EURO-96	1996	Log of match-day stadium attendance Log of stadium capacity utilization	1) Dummy indicating whether a match contains a seeded team	1) No	Also includes seasonal uncertainty

				2) Dummy indicating a match belongs to the knock-out stage	2) No	
Kuypers (1997)	England-1	1993/94	Match-day stadium attendance Proportion of Sky subscribers	1) Difference in maximum and minimum probabilities of home team win or draw and away team win	1) No	Also includes seasonal uncertainty
Falter and Perignon (2000)	France-1	1997/98 – 1998/99	Log of match-day stadium attendance	1) Home team rank 2) Away team rank 3) Goal-average differential between opponents	1) Yes* 2) No* 3) Yes	Also includes seasonal uncertainty, measured as season dummies summer, autumn and winter, spring being the reference, these dummies measure more than uncertainty
Czarnitzki and Stadtmann (2002)	Germany-1	1996/97 – 1997/98	Match-day stadium attendance	1) Home team rank 2) Away team rank 3) Home team win probability, also squared	1) Yes* 2) No* 3) No	Also includes seasonal uncertainty; formulate the Theil-index, but use home win probability
Forrest and Simmons (2002)	England-2/4	1997/98	Log of match-day stadium attendance	1) Ratio of home win probability and away win probability, also squared	1) Yes	Early season matches (in August and September) are excluded, as well all Premier League matches, correction for biases in bookmaker odds
García and Rodírguez (2002)	Spain-1	1992/93 – 1995/96	Log of match-day stadium attendance	1) Difference in rank measured home-away, also squared	1) No	Attendance is measured as tickets sold, excluding children and season tickets; also includes seasonal uncertainty
				2) Dummy indicating home team between +3 or -5 league positions of away team	2) No	
Forrest, Simmons and Buraimo (2005)	England-1	1993/94 – 2001/02	Log of TV audience	1) Difference in in-season points per game between opponents corrected for home advantage	1) Yes	First round of season is excluded; use pre- and post-Boxing Day period, also includes seasonal uncertainty
Forrest, Beaumont, Goddard and Simmons (2005)	England-2/4	1997/98	Log of match-day stadium attendance	1) Ratio of home win probability and away win probability. Controlling for home advantage	1) Yes	Early season matches (in August and September) are excluded, as well all Premier League matches; only matches played on Saturday are considered
				2) Points per game of the home team	2) No*	
				3) Points per game of the away team	3) No*	

Forrest and Simmons (2006)	England-2/4	1999/00 – 2001/02	Log of match-day stadium attendance	1) Difference in in-season points per game between opponents corrected for home advantage 2) Points per game of the home team 3) Points per game of the away team	1) No 2) No* 3) No*	The first three home fixtures for each team are excluded, separate results for the three divisions, also includes seasonal uncertainty
Bojke (2007)	England-2	2000/01	Match-day stadium attendance	1) Theil-index	1) Yes	Also includes seasonal uncertainty
Buraimo (2008)	England-2	1997/98 – 2003/04	Log of match-day stadium attendance Log of TV audience	1) Difference in in-season points per game between opponents corrected for home advantage 2) Points per game of the home team; only with stadium attendance 3) Points per game of the away team; only with stadium attendance	1) No 2) No* 3) No*	The first match of each season is excluded, furthermore, missing data on wages reduced the sample size
Benz, Brandes and Franck (2009)	Germany-1	1999/00 – 2003/04	Log of match-day stadium attendance	1) Difference in league standings 2) Difference in in-season points per game between opponents corrected for home advantage 3) Theil-index 4) Relative win probability based on Theil, without incorporation of the probability for a draw 5) Home team win probability, also squared	1) Partial 2) Partial 3) No 4) No 5) Partial	Stadium attendance figures are adjusted by subtraction of season tickets sold; observations with Bayern Munich being the away team and Derbies are excluded, because of too little variation in the dependent variable; partial support for UOH here indicates that it was found for high demand games only
Madalozzo and Villar (2009)	Brazil-1	2003-2006	Log of match-day stadium attendance	1) Home team rank 2) Away team rank 3) Difference in league standings	1) Yes* 2) Yes* 3) No	Due to various reasons, some matches are excluded, reducing the dataset form a potential of 1946 to 1851; attendance only contains paying visitors; also includes seasonal uncertainty

Pawlowski and Anders (2012)	Germany-1	2005/06	Log of match-day stadium attendance	1) Theil-index 2) Home favorite, dummy indicating if home team win probability is greater than away team win probability	1) No 2) No	Also includes seasonal uncertainty
Pawlowski (2013)	Germany-1	Survey in 2011/2012	Degree of CB (Competitive Balance)	1) Final outcome is unclear and home and away team have equal winning probabilities	1) Yes	Use of survey data; results suggest that UO matters, but improving competitive balance would not increase demand; also includes seasonal uncertainty
Buraimo and Simmons (2015)	England-1	2000/01 – 2007/08	Log of TV audience	1) Combined points per game 2) Difference in teams' probabilities of winning 3) Theil-index	1) No 2) Partial 3) Partial	Match uncertainty is also measured for single seasons; partial indicates that it is only significant for the first two seasons; also includes seasonal uncertainty
Serrano, García-Bernal, Fernández-Olmos and Espitia-Escuer (2015)	England, Germany, Italy, Spain-all 1	2012/13	Log of match-day stadium attendance	1) Theil-index, also squared	1) Partial	Quantile regression technique is used on a pooled dataset; partial support for UOH here indicates that it was found for high demand games only
Cox (2015)	England-1	2004/05 – 2011/12	Log of match-day stadium attendance Log of TV audience	1) Home win probability, also squared 2) Eight different dummies that represent home win probabilities 3) Probability of a draw 4) Absolute difference in home team win probability and away team win probability	1) No 2) No/Yes 3) No 4) No/Yes	Both 'yes' are based on the results for TV audience
Martins and Cró (2016)	Portugal-1	2010/11 – 2014/15	Log of match-day stadium attendance	1) Theil-index 2) Home favourite, dummy indicating if home team win probability is greater than away team win probability 3) Home win probability, also squared	1) No 2) No 3) No	Also includes seasonal uncertainty
Schreyer, Schmidt and Torgler(2016)	STH Borussia Dortmund	2012/13	Attendance and time of entrance	1) Theil-index, also squared 2) Home win probability, also squared	1) Yes 2) No	13,892 STH (season ticket holders)

				3) Home favorite, dummy indicating if home team win probability is greater than away team win probability	3) No	
				4) ROY; Theil-index without probability for a draw	4) Yes	
				5) Absolute difference in home team win probability and away team win probability	5) Yes	
				6) Absolute difference in league rank	6) Yes	
				7) Difference in in-season points per game between opponents corrected for home advantage	7) Yes	
Pawlowski, Nalbantis and Coates (2017)	Germany-1	Survey in 2014/15	Perceived game uncertainty	1) Subjective home win probability	1) No	Survey took place prior to the 10th and 27th Bundesliga matches; perceived game uncertainty coincides with match uncertainty measures commonly used

Note: * indicates this indicator actually measures team quality instead of uncertainty

Table C2: Overview of studies with seasonal uncertainty

Seasonal uncertainty	Country-Division	Seasons	Dependent Variable	Seasonal Uncertainty	UOH	Comment
Jennett (1984)	Scotland-1	1975/76 – 1980/81	Log of match-day stadium attendance	1) Championship significance home team, still able to win, increasing throughout season 2) Championship significance away team, still able to win, increasing throughout season 3) Relegation significance, still able to avoid, increasing throughout season	1) Yes 2) Yes 3) No	Jennett describes these significance variables as measures of short term uncertainty
Crains (1987)	Scotland-1	1971/72 – 1979/80	Match-day stadium attendance	1) Championship contention measured by dummy variable, still able to win	1) Yes	Separate analyses for a selection of three teams; also included match uncertainty
Janssens and Késenne (1987)	Belgium-1	1982/83	Match-day stadium attendance	1) Championship significance home team; own measure 2) Championship significance away team; own measure	1) Yes 2) No	Also includes match uncertainty; measure of seasonal uncertainty is a variant to Jennett's measure
Dobson and Goddard (1992)	England-1/4	1989/90 – 1990/91	Log of match-day stadium attendance, distinguishing standing and seating	1) Championship/promotion significance home team, still able to win, increasing throughout season (logarithm) 2) Championship/promotion significance away team, still able to win, increasing throughout season (logarithm)	1) Yes 2) No	Analyses based on a set of clubs that (were able to) provide survey information; also includes match uncertainty; Jennett's measures adapted
Wilson and Sim (1995)	Malysia-1/2	1989/90 – 1991/92	Log of match-day stadium attendance	1) Championship significance home team, still able to win, increasing throughout season 2) Championship significance away team, still able to win, increasing throughout season	1) Yes 2) No	Non-paying spectators and season ticket holders are excluded from the analyses; also includes match uncertainty; Jennett's measure
Baimbridge, Cameron and Dawson (1996)	England-1	1993/94	Log of match-day stadium attendance	1) Championship significance indicated by dummy if both teams are ranked in the top four 2) Relegation significance indicated by dummy if both teams are ranked in the bottom four 3) Home match trend, also squared	1) No 2) No 3) Partial	Also includes match uncertainty; the trend variables are argued to capture seasonal uncertainty, but also some other aspects such as seasonal weather conditions
Baimbridge (1997)	EURO 96	1996	Log of match-day stadium attendance Log of stadium capacity utilization	1) Match significance measured by the mean of winning probabilities 2) Match trend, also squared	1) Yes 2) Partial	Also includes match uncertainty
Kuypers (1997)	England-1	1993/94	Match-day stadium attendance Proportion of Sky subscribers	1) Championship significance; measured by own indicator 2) Relegation significance; measured by own indicator	1) Yes 2) No	Also includes match uncertainty; measure of seasonal uncertainty includes components of the number of points

						behind the leader and the remaining number of matches
Falter and Perignon (2000)	France-1	1997/98 – 1998/99	Log of match-day stadium attendance	1) Seasons of the year dummies	1) Yes	Also includes match uncertainty; seasonal uncertainty measured as dummies for summer, autumn and winter, spring being the reference, these dummies measure more than uncertainty
Czarnitzki and Stadtmann (2002)	Germany-1	1996/97 – 1997/98	Match-day stadium attendance	1) Championship significance home team, still able to win 2) Championship significance away team, still able to win	1) No 2) No	Also includes match uncertainty; seasonal uncertainty as measured by Janssens and Késenne (1987)
García and Rodírguez (2002)	Spain-1	1992/93 – 1995/96	Log of match-day stadium attendance	1) Championship significance of home team, measured as the product of the number of points behind the leader and the number of games left before championship is decided, being zero in case the championship is decided	1) Yes	Attendance is measured as tickets sold, excluding children and season tickets; also includes match uncertainty; seasonal uncertainty is measured by an indicator of Kuypers (1997)
Forrest, Simmons and Buraimo (2005)	England-1	1993/94 – 2001/02	Log of TV audience	1) Set of dummies indicating whether a match is important for some end-of-season achievement	1) Yes	First round of season is excluded; use pre- and post-Boxing Day period, also includes match uncertainty
Forrest and Simmons (2006)	England-2/4	1999/00 – 2001/02	Log of match-day stadium attendance	1) Promotion contention dummies interacted with month dummies	1) Yes	The first three home fixtures for each team are excluded, separate results for the three divisions, this measure is not documented as seasonal uncertainty by the authors, also includes match uncertainty
Bojke (2007)	England-2	2000/01	Match-day stadium attendance	1) Promotion probability; own measure	1) Yes	Also includes match uncertainty; own measure of promotion probabilities simulated with match level bookmaker odds
Madalozzo and Villar (2009)	Brazil-1	2003-2006	Log of match-day stadium attendance	1) Chance of being league leader 2) Chance of going to the Libertadores Cup 3) Chance of leaving relegation zone 4) Game position in schedule	1) Yes 2) No 3) Yes 4) Yes	Due to various reasons, some matches are excluded, reducing the dataset form a potential of 1946 to 1851; attendance only contains paying visitors; also includes match uncertainty
Pawlowski and Anders (2012)	Germany-1	2005/06	Log of match-day stadium attendance	1) Championship significance home team, still able to win, increasing throughout season	1) Yes	Also includes match uncertainty; seasonal uncertainty as measured by Jannens and Késenne (1987)

				2) Championship significance away team, still able to win, increasing throughout season 3) Champions League significance home team, still able to qualify, increasing throughout season 4) Champions League significance away team, still able to qualify, increasing throughout season	2) Yes 3) No 4) Partial	
Pawlowski (2013)	Germany-1	Survey in 2011/2012	Degree of CB (Competitive Balance)	1) Championship, Champions League, Europa League and relegation contentions are exiting	1) Yes	Use of survey data; results suggest that UO matters, but improving balance would not increase demand; also includes match uncertainty
Buraimo and Simmons (2015)	England-1	2000/01 – 2007/08	Log of TV audience	1) Championship contention, good opportunity to win 2) Contention for qualification for European football, good opportunity to qualify 3) Relegation contention	1) No 2) No 3) No	Contention is measured as: dummy indicating that in order to achieve result, all remaining games are won and all other teams draw, for relegation if all other teams win and the selected team draw; also includes match uncertainty
Martins and Cró (2016)	Portugal-1	2010/11 – 2014/15	Log of match-day stadium attendance	1) Championship significance home team, still able to win, increasing throughout season 2) Championship significance away team, still able to win, increasing throughout season 3) Champions League significance home team, still able to qualify, increasing throughout season 4) Champions League significance away team, still able to qualify, increasing throughout season 5) Cumulative changes in rank order 6) Rank order changes 7) Standard deviation on weekly changes in winning percentages	1) Yes 2) No 3) Yes 4) No 5) No 6) No 7) No	Also includes match uncertainty

Appendix D: Alternative baseline result

Table D1: Baseline results with Theil and PPG

	(1)	(2)	(3)	(4)	(5)	(6)
Derby	0.043***	0.043***	0.041***	0.044***	0.046***	0.041***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Weekday	-0.028***	-0.029***	-0.025***	-0.027***	-0.028***	-0.024***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Temperature	-0.008***	-0.009***	-0.008***	-0.005*	-0.009***	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Precipitation	-0.010***	-0.009***	-0.010***	-0.009**	-0.009***	-0.009***
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
Odd-Surprise Team/100	0.281***	0.232***	0.250***	0.377***	0.284***	0.297***
	(0.076)	(0.073)	(0.073)	(0.077)	(0.075)	(0.074)
Odd-Surprise Opponent/100		0.082**	0.084***		0.033	0.045
		(0.033)	(0.030)		(0.035)	(0.032)
Quality Team/100	0.059	0.270*	0.248*	0.068	0.265*	0.238*
	(0.147)	(0.138)	(0.136)	(0.143)	(0.139)	(0.136)
Quality Opponent/100		0.303***	0.077***		0.284***	0.055**
		(0.017)	(0.027)		(0.016)	(0.026)
Theil	0.010	-0.096***	-0.038**			
	(0.023)	(0.017)	(0.018)			
PPG				-0.026***	-0.004	-0.008***
				(0.003)	(0.003)	(0.003)
Championship	0.037	0.018	0.008	0.036	0.011	0.005
	(0.044)	(0.042)	(0.041)	(0.043)	(0.043)	(0.041)
UEFA Champions League	0.059	0.054	0.067*	0.054	0.057	0.067*
	(0.041)	(0.040)	(0.039)	(0.041)	(0.041)	(0.039)
UEFA Europa League	0.102***	0.087***	0.089***	0.089***	0.084***	0.085***
	(0.023)	(0.025)	(0.023)	(0.023)	(0.025)	(0.023)
Relegation	0.161***	0.171***	0.165***	0.158***	0.170***	0.163***
	(0.040)	(0.042)	(0.041)	(0.039)	(0.041)	(0.041)
Opponent FE	No	No	Yes	No	No	Yes

Note: Tobit regression with attendance rate as dependent variable and with the upper limit set at 0.95. All estimates in models (1)-(3) contain 4,586 observations (1,890 censored), all estimates in models (4)-(6) contain 4,588 observations (1,891 censored). All models contain 270 club-season fixed effects and robust standard errors in parentheses, clustered by club-season. *Odd-Surprise Team, Odd-Surprise Opponent, Quality Team* and *Quality Opponent* are divided by 100; ****p<0.01, **p<0.05, *p<0.1.*

Chapter 4

Selection of top-level talent and the relative age effect: a study on elite youth football players

4.1 Introduction

The labour market for top-level talents is characterized by a high degree of competition between the candidates. On the one hand, positions where highly talented people are needed are scarce. On the other hand, it is reasonable to assume that many people want to be in such a position, since rewards are high. In business, this holds for the position of, for example, a CEO. In sports, one can think of professional players. Since it is in the interest of organizations to find the most talented people for these key positions, it is important that the selection system for talent functions properly. However, evidence exists that selection of talent can be far from optimal. One phenomenon that points at a distortion of the selection system for talent is the Relative Age Effect (RAE). Relative age differences between people arise, because they are grouped together based on birth dates and groups, in general, cover a full year. Furthermore, the groups are separated by a cut-off date such as in school grades and sports teams. The RAE implies, that the relative age differences between people result in the overrepresentation of peers that are born close to the cut-off date and, therefore, experience a relative age-advantage. Someone who has been born just posterior to the cut-off date is early-born and relatively old compared to someone in the same group, who has been born just prior to the cut-off date. Especially at young ages, the relative age differences can be quite substantial. In many situations, there is an advantage of being relatively old, for example, because of superior physical and cognitive development. This advantage generally results in superior current performances. If selection systems for talent are unable to incorporate good predictions of future or potential performances, the focus lies on these current performances. Then, an overrepresentation of the age-advantaged peers (i.e. an RAE) is likely to occur.

In sports, such an overrepresentation could be the result of rational choices made by coaches. They are often former (professional) players themselves and motivated and trained to win matches. Furthermore, extensions of employment contracts may depend on team performances. This serves as an external motivational factor for winning. It seems reasonable that winning can best be achieved by picking the best current players, who are not necessarily the most talented ones. Assuming that coaches also care about the development of talent, they experience a conflict in goals. As far as winning is the dominant driver, the occurrence of the RAE may be the result of rational choices. However, club officials generally state that talent selection and development matter most within youth academies and should have the focus of coaches. Still, in the press, journalists and pundits often use the focus on winning matches as an argument why youth talent does not develop as it should be.

A few problems are related to relative age effects. First, talent is generally assumed to be uniformly distributed across birth dates. An overrepresentation of age-advantaged people means that some less talented persons are selected in favour of more talented, but age-disadvantaged, peers. Furthermore, many selection systems already start at young ages, when relative age differences are large. This is typically the case in sports. Consequently, the age-advantaged peers that have been selected, are provided with better facilities and support to practice and to develop skills. For example, by participation in talent development programs in sports, or specifically in football, by participation in the youth academies of professional clubs. Since the age-disadvantaged, but equally talented, people are not selected, they will never receive equal opportunities to develop, compared to their age-advantaged peers. No extra support is given and, as a consequence, they might never reach superior performance levels at older ages. However, this might have been possible with such extra support and given their initial talent. Thus, this talent is lost, because the selection system is not able to correctly incorporate relative age differences.

Many previous studies have looked at the RAE. This literature will be discussed in the next sections. In general, it is found to be persistent across different domains, such as school systems and sports. The area of sports is popular among scholars, because data is rather easily available and rules and regulations regarding group compositions and cut-off dates are clearly defined. Furthermore, since professional sports is an interesting industry in itself, it is important to understand how talent selection works and how this can be improved.

In this study, I use data from elite youth football. More specifically, I use data from the youth academy of PSV Eindhoven (PSV), which is one of the top clubs in Dutch professional football. The main aim of the academy is to develop individual talent, as can be found in the mission statement of the youth academy.[1] The dataset contains all players who have been born in 1988 or later and were active in the academy for at least one season. In Dutch football, the relevant cut-off date is January 1, thus children are grouped together by calendar year. First, I use some descriptive statistics and a χ^2-Goodness-of-fit-test to show the persistence of the overrepresentation of players who are born in the first part of the year, i.e. the presence of an RAE. I use birth quarters as well as birth semesters for

[1] http://www.psv.nl/jeugd/opleiding/missie.htm; Dutch only. Retrieved November 2017.

this. Then, I make a distinction between external selection and internal selection. External selection deals with the recruitment of players from outside of the academy. Internal selection is focused on the annual decision whether players can stay within the academy or have to leave the academy. Since I have data on the full playing histories within the academy of all players, I can study these decisions quite accurately. The main difference between the types of selection, is the intensity with which players are observed prior to the decision to select. External selection is based on scouting reports and a rather low number of observations. In particular, these observations include current performances. Internal selection is based on a high number of (almost daily) observations. This high number of observations reduces the variation that makes the prediction of (potential) talent more accurate. In general, one may assume that external selection contributes to the RAE, while internal selection does not contribute, or even reduces the severity of the RAE. Descriptive statistics are in line with both assumptions. Furthermore, a linear probability model reveals that early-born players have a higher probability to leave the academy, compared to their late-born peers. However, the reduction of the number of these early-born players that results from this internal selection, is not large enough to overcome the biased birth-date distribution that results from external selection. Finally, I investigate whether there is a difference in the probability of becoming a professional football player between early-born players and their late-born peers. Based on a linear probability model, I find that, for players aged 19, the late-born players have a higher probability to become a professional. Furthermore, it also follows that the majority of professionals is born in the first part of the year. Combined, these results suggest that the selection system only picks the highly talented late-born players. It also suggests that some late-born players are not selected, although they are more talented than some of their early-born peers, who have been selected.

Although a study on the RAE within elite youth football is not necessarily new, the dataset that I use is rather unique. It allows me to investigate the persistence of the RAE within the youth academy of a top club within the Netherlands. Furthermore, since the playing histories within the academy are included, it makes the distinction between external and internal selection possible. The results for internal selection provide new insights, which have not been discussed by previous studies. Finally, my analysis on the difference in probability to become a professional football player between early-born players and late-born peers, reveals the consequence of the biased selection system in elite youth football. This adds to the discussion on the loss of talent.

The rest of this paper is organized as follows. Section 4.2 provides a brief discussion on talent selection, identification and development. Next, in section 4.3, some previous literature on the RAE is discussed, while section 4.4 describes the data that is used. Section 4.5 presents the results on the

RAE. Section 4.6 shows the results on the difference between external and internal selection mechanisms. In section 4.7, the relation between birth semester and the probability to become a professional football player is discussed. Section 4.8 presents a discussion, with separate subsections on some assumptions made and on the generalizability of results. Finally, section 4.9 concludes.

4.2 Talent selection, identification and development

Talent selection, identification and development covers a wide range of aspects. Rees *et al.* (2016) provide a recent overview of almost 300 studies in which they distinguish between three main elements. These are *the performer*, *the environment* and *practice and training*. For each element, several topics are discussed. One of these topics is *birthdate*, which relates to the RAE. Their recommendation for practitioners and policy makers is that RAEs should not be used in talent selection. Furthermore, the environment should be structured such that the negative effects of relative age are limited. Overall, based on their evaluation of all topics, they conclude that many aspects matter for talent selection, identification and development, probably in an interactive way. In this section, I will briefly discuss only a limited number of these aspects and studies. In doing so, I follow the development over time of the way in which scholars have been thinking about talent. For a comprehensive discussion, I refer to an academic review by Abbott *et al.* (2002) and the review article by Rees *et al.* (2016).

For a long time, talent was assumed to be innate and genetically determined. This nature-based view suggests that nurture (i.e. things like practice and the environment) is of little importance for obtaining expertise. This changed after the seminal work by Ericsson, Krampe and Tesch-Römer (1993) on the role of deliberate practice (DP) for the acquisition of elite levels of performance. Deliberate practice is effortful, but not necessarily enjoyable. Someone should be motivated to conduct a lot this practice to become an expert. Based on data for musicians, the authors show that the amount of DP determines whether someone becomes an elite performer. In general, it will take about ten years of such practice to achieve expert levels (Ericsson, Krampe and Tesch-Römer, 1993). Since this coincides with about 10,000 hours, this has become known as the '10,000 hour rule', which, amongst others, is also found for chess (Simon and Chase, 1973) and is referred to in more popular books as well (e.g. Gladwell, 2008). In line with the results of Ericsson, Krampe and Tesch-Römer (1993), Howe, Davidson and Sloboda (1998) question the role of innate talent and argue that early signs of expertise result from early experiences and opportunities to learn. Furthermore, they conclude that a sufficient amount of practice is needed to achieve expert performance levels.

There is evidence that suggests that genetics play a role in obtaining elite performances (Rees *et al.*, 2016). However, nowadays, most people are convinced of the importance of practice as well. Therefore, scholars have been applying the concept of deliberate practice to sports. For example, Ward *et al.* (2007) investigate the role of DP in English youth football, while Helsen *et al.* (2000) specifically discuss the role of DP in youth football, acknowledging that an RAE is present there. In general, the results are similar to those of Ericsson, Krampe and Tesch-Römer (1993). However, in contrast, practice in sports, and more specifically, in team sports, is often enjoyed by the participants. Furthermore, the development of abilities and skills may take place during playful activities that are not necessarily structured as practice (Rees *et al.*, 2016). Over time, it has been argued that several additional elements are relevant within the framework of DP. For example, Cumming and Hall (2002) discuss the role of deliberate imagery practice and find this to be important for elite performances.

Although the DP framework has appealing elements and deliberate practice is found to be relevant for expertise, some concerns are raised. The framework assumes domain-specific and monotonic benefits. Simonton (1999), amongst others, suggests that this is not necessarily the case and formulates a model in which (innate) talent is multidimensional, multiplicative and dynamic. Although he acknowledges the importance of environmental factors and deliberate practice, his model assumes that talent is also relevant for the development of expertise. In contrast to the DP framework, the dynamic component in Simonton's model allows for early bloomers and late bloomers, as well as for the loss of talent. Furthermore, since talent develops in a person- and domain-specific way, an individual's optimal domain may change over time. In that respect, scholars have started to think of talent in a more multidimensional, multiplicative and dynamic way. Examples include Williams and Reilly (2000) and Abbott and Collins (2004). Williams and Reilly (2000) formulate a theoretical model for talent identification and development in (male) football. They distinguish between talent *detection* (of potential performers), talent *identification* (of elite performers), talent *development* (via suitable opportunities and learning environments) and talent *selection* (which is an ongoing process of picking the best players). Furthermore, the authors argue that multiple elements are relevant for the prediction of talent, which include sociological, physical, physiological and psychological aspects. Abbott and Collins (2004) stress the importance of the psychological elements. They suggest that talent identification and development should be focused on an individual's progression instead of early identification that is based on a low number of observations (as is the case with scouting in football). If sufficient opportunities are provided, motivation and self-regulatory learning strategies will guide the development of performances (Abbott and Collins, 2004). Empirical evidence from elite Dutch youth football supports the idea that self-regulatory strategies and performance levels are positively related (Toering *et al*., 2009). Furthermore, Toering *et al*. (2011) use relative age as a moderator variable, which does not affect the positive relation and, therefore, may serve as a useful measure in talent identification.

Although the previous paragraph shows that, over time, talent has been viewed as dynamic and multidimensional, Vaeyens *et al*. (2008) argue that a theoretical framework was missing. They criticize cross-sectional models of talent identification and development, since these do not correctly incorporate the dynamic and multidimensional aspects. A skewed birth-date distribution and the missing of late bloomers are problems that result from these models. Improvements could be made through a distinction between potential and performance, while accounting for all possible determinants of talent. Vaeyens *et al*. (2008) suggest that the Differentiated Model of Giftedness and Talent (DMGT) may be useful in that respect. This model was initially formulated by Gagné (1993) in the domain of education and has been adapted multiple times (e.g. Gagné, 2004; Gagné, 2010). In general, the model distinguishes between natural gifts (giftedness) and an end-product of development (talent) and is used as such in a sports context by Vaeyens *et al*. (2008). Many aspects are included in an interactive way, so that it accounts for the multidimensional and dynamic elements of talent. As such, it allows for the evaluation of progression instead of performance. A full discussion of the model is beyond the scope of this paper. My conclusion is that it seems to be a useful theoretical tool to think about talent identification and talent development.

4.3 Related literature on the RAE

As mentioned in the introduction, many different topics have been related to RAEs. These include school systems and academic achievements (e.g. Angrist and Krueger, 1991; Plug, 2001) and high school leadership (Dhuey and Lipscomb, 2008). Furthermore, the RAE is studied in relation to the development of self-esteem (Thompson, Barnsley and Battle, 2004) and youth suicide rates (Thompson, Barnsley and Dyck, 1999). The focus in this section will be on previous studies that investigate the RAE in a sports context. An overview is given by Musch and Grondin (2001), while Cobley *et al.* (2009) provide a meta-analytical review. Furthermore, Wattie, Schorer and Baker (2015) review previous studies and propose a constrained-based theoretical framework that relates to the RAE. In general, a skewed birth-date distribution is found in favour of the relative age-advantaged peers. This result is not necessarily the same for both genders, but I neglect this difference in this section and do not separately discuss males and females.

The first to study the RAE in sports are Grondin *et al.* (1984) and Barnsley, Thompson and Barnsley (1985). Grondin *et al.* (1984) find an overrepresentation of age-advantaged players for different levels of Canadian ice hockey. For volleyball, this is only visible for the elite level. Barnsley, Thompson and Barnsley (1985) only investigate the elite adolescent and elite senior levels of Canadian ice hockey and find and overrepresentation of the early-born players. Furthermore, in a follow-up study, Barnsley and Thompson (1988) reveal that this is also the case for junior divisions. After these initial studies, many contributions followed that use data on ice hockey and find an RAE (e.g. Sherar *et al.*, 2007; Addona and Yates, 2010 and Nolan and Howell, 2010). Furthermore, an RAE is found in baseball (Barnsley, Thompson and Stebelsky, 1991), basketball (Hoare, 2000), swimming and tennis (Baxter-Jones *et al.*, 1995), rugby (McCarthy and Collins, 2014; McCarthy, Collins and Court, 2016), cricket (McCarthy, Collins and Court, 2016) and in combat sports (Albuquerque *et al.*, 2016). However, no RAE was found for gymnastics (Baxter-Jones *et al.*, 1995) and dancing (Van Rossum, 2006), while an RAE is only found for a few male age categories in the shooting sport (Delorme and Raspaud, 2009). Of main interest to the present study, is the vast amount of research that has been conducted in football (e.g. Verhulst, 1992; Dudink, 1994; Baxter-Jones *et al.*, 1995; Helsen, Starkes and Van Winckel, 1998; Musch and Hay, 1999; Helsen, Starks and Van Winckel, 2000; Helsen *et al.*, 2000; Simmons and Paull, 2001; Helsen, Van Winckel and Williams, 2005; Vaeyens, Philippaerts and Malina, 2005; Vincent and Glamser, 2006; Ashworth and Heyndels, 2007; Jiménez and Pain, 2008; Delorme, Boiché and Raspaud, 2010; Augste and Lames, 2011; Krikendall, 2014; Doyle, Bottomley and Angell, 2017; Mann and Van Ginneken, 2017). An RAE is generally found to be present. Furthermore, Helsen *et al.* (2012) document that ten years of research have not changed the occurrence or severity of the RAE in European professional football.

Besides studying the existence of RAEs, scholars have been interested in the reasons and mechanisms for it to arise. In general, the organisational structure of competitive sports and the use of a cut-off date for age groups, are seen as main elements in that respect. There is direct evidence that suggests that the cut-off date indeed is important for the RAE to occur. This evidence is based on the shift in cut-off date in 1997 in football. Musch and Hay (1999) show, for Australian football, that the overrepresentation of players changes in line with the change in cut-off date, from August 1 to January 1. A similar result is found for Belgian football by Helsen, Starks and Van Winckel (2000) and Helsen *et al.* (2000). Furthermore, Simmons and Paull (2001) show that the difference in birth-date distributions between samples for English and UEFA youth selections, corresponds with the difference in cut-off date. For the sample of English players, this is September 1, while for the sample of UEFA selections, this is January 1.

Although the cut-off date is an important organisational element for the occurrence of the RAE, at least within football, it has no direct effect on the selection of players. Therefore, it seems likely that the grouping of players by age that results from the use of a cut-off date, provides the early-born players with an advantage. In particular, this advantage might be a physical advantage that early-born players have over their relatively younger peers (e.g. Musch and Grondin, 2001; Cobley *et al.* 2009). Sherar *et al.* (2007) provide evidence, based on Canadian junior ice hockey selections, in support of this assumption. Other, more indirect evidence of the importance of physical development, comes from sports in which such development is not necessarily an advantage. For example, Baxter-Jones *et al.* (1995) do not find an RAE in gymnastics. Furthermore, Van Rossum (2006) does not find an RAE for dancing and argues that this is not surprising, since technical skills are more important than physical development. Finally, Delorme and Raspaud (2009) only find an RAE for a few male age categories in the French shooting sport. They argue that in this sport, concentration is more important than physical development. Although direct evidence is missing for football, it is comparable to ice hockey in terms of physical intensity. Thus, the above suggests that, within football, physical development (i.e. early maturation) results in an advantage at early ages. Over time, this physical advantage will reduce, since almost everyone will mature. For that reason, one may expect a decrease of the overrepresentation of age-advantaged players over time. However, Cobley *et al.* (2009), for example, find an increase up to the ages of adolescence, while a decrease at adult ages. The increase at junior ages is in line with higher drop-out rates of late-born players in football (Helsen, Starkes and Van Winckel, 1998). For the decrease of the overrepresentation at adult ages, Cobley *et al.* (2009) give three potential reasons. First, players may change their sport. Second, early-born players will resign from the sport at an earlier age, because of an injury as a result of overtraining in talent development programs. Third, as discussed, the physical advantage of early-born players may disappear.

A further element that is related to the occurrence of an RAE is the amount of competition within a given sport. Musch and Grondin (2001) argue that RAEs are expected to be larger within sports where competition is large. With a simple example they explain that, if the number of available spots in a team is equal to the number of applicants, no selection takes place and, thus, there is no reason for an RAE to occur. In general, the amount of competition is measured by playing level (i.e. non-elite vs. elite) and the popularity of the sport. Since football is very popular (at least in Europe) and given the severe RAEs that are found within football, Musch and Grondin (2001) see this as evidence in favour of the assumption that the degree of competition matters.

So far, the discussion of previous studies reveals that talent selection and the organisational structure of competitive sports, favours age-advantaged players. In addition, Augste and Lames (2011) suggest that playing with an, on average, older team, is beneficial for team success in elite German youth football. However, their results reveal a rather high amount of variation and heterogeneity between teams, from which they conclude that a high relative age is not a necessity to compete in the top of these youth competitions. Furthermore, Kirkendall (2014) suggests that the RAE has no influence on match outcomes in US youth football. This means that there is no reason not to play with late-born players. Although still limited, there is also evidence of a reversal of the RAE, i.e. late-born players are better off than their early-born peers at adult ages. Recent examples are McCarty and Collins (2014) for rugby and McCarty, Collins and Court (2016) for rugby and cricket. Both studies find an overrepresentation of early-born players at elite junior levels. However, at senior ages, the conversion rate into professionalism is higher for the age-disadvantaged players. Moreover, Ashworth and Heyndels (2007) find that the salaries within German Bundesliga football are higher for the late-born players. Although Furley, Memmert and Weigelt (2016) find that the market values are higher for early-born players (based on a sample of the 100 most valuable football players according to CIES 2015) there seem to be benefits for the age-disadvantaged players who enter the talent selection and development process at junior levels. Of course, this may reflect their initial high potential and talent. Without this, they may never have been selected in the first place.

Despite the positive outcome for the few highly talented late-born players, the RAE generally results in a waste of talent. Doyle, Bottomley and Angell (2017) show that the waste of talent is substantial and they estimate it to be 57 percent for English Premier League youth teams. Thus, this suggests that there is plenty of reason to think of remedies to this problem. Musch and Grondin (2001) discuss the possibility of classifications that are based on biological age, such as weights. Furthermore, they propose a classification that is based on chronical age and a rotation of the relative age advantage. This may result from an annual shift in the cut-off date, by, for example, three months. A different solution they discuss is to use multiple squads that are based on multiple standards. In this situation, players may stay on a lower level if they do not yet meet certain criteria, while their age prescribes to continue. Finally, they point out the importance to warn practitioners for the RAE. Except for this last suggestion, which seems to be a good idea anyway, these solutions are rather difficult to implement, especially for team sports. Furthermore, they may have unwanted consequences outside of the sports. For example, at a young age, it may be better for the development of children not to separate them from their friends. Implementing age quotas is an easy solution to the RAE as well (e.g. Barnsley and Thompson, 1988) but is never tested and contradicts with the idea that experts (i.e. scouts and coaches) are able to detect potential and talent. In that respect, a recent experiment by Mann and Van Ginneken (2017) provides some interesting results. They show that scouts are able to incorporate relative age differences in their evaluations of playing talent, if the information on relative age is presented to them appropriately. No improvement was found when birth dates of players were given. However, in the situation that the shirt numbering represented the relative age differences between players, the RAE was reduced. Of particular interest is that the scouts that participated in the experiment, were all active at PSV. Out of 25 scouts, 22 said to be familiar with the RAE prior to the experiment (Mann and Van Ginneken, 2017). Furthermore, Art Langeler, the head of the youth academy between July 2013 and June 2017, said in an interview with a local newspaper[2] that PSV is familiar with the problems that arise as a result of the RAE. Therefore, they specifically take it into account within their selection system. Thus, PSV seems interested in improvements of their talent selection. In the rest of this paper, we will see how well the scouts did over the past two decades. Given that the RAE and the consequence of a loss of talent is a known phenomenon, one would think that the RAE is small within the academy. However, although familiarity exists, their might always be an implicit focus on current performances.

4.4 Data

The data comes from the youth academy of PSV and is available on the internet via www.psv.nl/jeugd.[3] The youth archive contains all players who have been active in the youth academy as of the 1996/97 season up to and including the 2016/17 season. In total, the dataset contains 860 players. Since information for players who were active in the earlier seasons is incomplete, in particular concerning the time that someone entered the academy, I will only consider the players who were born in 1988 or later. In doing so, I leave out all selections that might have been influenced by the shift in cut-off date in the season 1997.[4] In total, the dataset contains 553 player-observations.[5] For all these players, a birthdate is available as well as their whole playing history within the youth academy of PSV. Thus, I have data on the time that someone entered the academy, when he left the academy (if applicable) and, for each season, in which team he played. Furthermore, I know the nationality of the player and whether someone is a goalkeeper (unknown for 6 players who are born in 1988).

[2] Eindhovens Dagblad, 15 July 2015; in Dutch.

[3] In fact, all data used to be available on www.psvjeugd.nl until the beginning of 2017. Since the data was collected prior to the change in website, this actually is the main source. Although some differences exist, especially in the way things are presented, most information has remained unchanged.

[4] During that season, the cut-off date changed from August 1 to January 1, as described in more detail in the previous section.

[5] One player left the academy at some point in time, but returned a few years later. This player is considered as a new player-observation when he returned.

The rules regarding age categories and the corresponding cut-off dates are set by the FIFA and the KNVB and prescribe in which team someone is allowed to play. Someone is not allowed to play in a lower age category, i.e. with younger players, while there is no problem to play in a higher age category, i.e. with older peers.[6] In general, players can be active within the youth academy between ages eight and 19. However, this differs between seasons, depending on the number of teams and age categories that are covered within the academy. Especially for the younger age groups, some changes were made in the number of teams throughout the sample period. Over time, some teams were added to the academy for players aged nine and 10. For the older age categories, some changes were made as well. For example, in some seasons there were two teams with players aged between 15 and 17, while in other seasons there was only one team. Naturally, these changes result in differences in the number of players that enter and leave the academy. However, these differences do not affect the analysis.

Table 4.1: Overview of heads of the youth academy of PSV

Name	Nationality	Start	End
Huub Stevens	Dutch	July 1986	March 1993
Hans Westerhof	Dutch	July 1993	June 1994
Frank Arnesen	Danish	July 1994	June 1996
Tonny Bruins Slot	Dutch	July 1996	June 1998
Remy Reijnierse	Dutch	July 1998	June 2001
Fred Rutten	Dutch	July 2001	June 2002
Pim Verbeek	Dutch	July 2002	June 2003
Joop Brand	Dutch	July 2003	June 2006
Edward Sturing	Dutch	July 2006	June 2008
Wiljan Vloet	Dutch	July 2008	November 2009
Jelle Goes	Dutch	December 2009	September 2012
Art Langeler	Dutch	July 2013	June 2017
Pascal Jansen	Dutch	July 2017	

Over time, several people have been in charge of the youth academy of PSV. An overview is given in Table 4.1. Most of them are former Dutch professional football players. In periods without an official head, the tasks were taken over by the management of the club. It seems reasonable that all had their own policy and preference for a certain organizational structure. However, I lack information on changes in policy during the sample period. In general, the youth academy has become more important for the club over time, given her ambitions and the financial differences between the Netherlands and other Europeans countries. Therefore, PSV states that she improved their youth

[6] Age categories typically cover two calendar years and used to be denoted with the letters A (age 18 and 19; oldest group) through F (age eight and nine; youngest group). Furthermore, it is common to distinguish between first year players who play in a second team (e.g. C2) and second year players who play in a first team (e.g. C1). Currently, the categories are denoted with O19 (under nineteen) through O9 (under nine) and a difference can be made between every single calendar year.

policy as of July 2013, with a specific goal to increase the number of players that make it into the first team.[7] In relation to the present study, I will provide some statistics in the next section that suggest that any policy change did not affect the persistence of the RAE within the academy.

Table 4.2 provides an overview of the number of players within the academy by team, age category and age. It follows that most players are active within their correct team, i.e. in the team that corresponds with their age. Note that these ages indicate that a player becomes this age in a given calendar year. Thus, the age of 16 indicates that the player turns 16 in the calendar year and is not aged like that at the start of the year. Only in three situations, a player was active in a lower team. Two times, a player was selected for the D2 instead of the D1, while once a player was active within the C2 instead of the C1. Furthermore, it appears to be rather uncommon to play in an older team, while it is even more uncommon for junior players to play for one of the senior teams. Finally, on a few occasions a player went on loan to a another club. Note that a single player is counted for each season that he is active within the academy. In total, the 553 player-observations, result in 2,410 player-team observations. Since the players who are promoted to one of the senior teams, in fact leave the academy, I exclude the 49 observations for Young PSV and First team PSV in (most of) the analysis.

Table 4.2: Distribution of players between team and age

Age category	F		E		D		C		B		A		
Team/Age	8	9	10	11	12	13	14	15	16	17	18	19	Total
Younger team						2		1					3
Correct team		140	211	259	282	274	258	237	195	144	126	84	2,210
Older team	14	7	10	14	12	10	12	8	13	29		1	130
Young PSV										1	8	28	37
First team PSV											4	8	12
On Loan			2			3	1	3	7	1		1	18
Total	14	147	223	273	294	289	271	249	215	175	138	122	2,410

Note: Most players play in their correct team, i.e. the team that corresponds with their age. Younger team means that someone is playing in a lower that than would be expected given his age (e.g. C2 instead of C1). Older team means that someone is paying in an older-age team within the academy. Youth means that someone has an age to play in the academy, but is part of 'Young PSV' (senior). First means that someone has an age to play in the academy, but is part of the First team of PSV (senior). On loan means that someone is on loan to another club.

Leaving out these observations results in 530 players. Table 4.3 shows to which age cohort (i.e. birth year) these players belong. In total, 21 different age cohorts are covered. The final column gives, for each cohort, the total number of players that have been active in the academy. Thus, 30 in the first row indicates that 30 different players, all born in 1988, have been active for at least one season in the academy. Furthermore, the table provides information, per age cohort, on the number of players

[7] Mission statement on the website of the youth academy (http://www.psv.nl/jeugd/opleiding/missie.htm; Dutch only). Retrieved November 2017.

that were playing for PSV at a certain age. For example, for the 1988 cohort, the value 13 for age 17 indicates that, at this age, there were 13 players active in the academy that were born in 1988. It follows that the majority of observations concentrates around the ages 11 up to and including the age of 15. This results from the number of teams within these age groups, since for (almost all cohorts of) these age groups, a single team was available. In comparison, for the age category A only one team is available. Finally, Table 4.3 shows, given a certain birth year, the final age group that is included in the sample. Thus, for cohort 1999, no players have reached the age (group) of 19 yet.

Table 4.3: Distribution of players between cohort and age

Age category	F		E		D		C		B		A		
Cohort/Age	8	9	10	11	12	13	14	15	16	17	18	19	N. Player
1988			2	7	18	18	16	17	18	13	9	7	30
1989				15	19	17	17	19	13	13	7	8	33
1990			2	13	14	18	18	16	15	11	8	6	34
1991			1	13	13	16	16	16	12	10	8	8	26
1992			10	15	17	16	17	13	14	12	8	8	29
1993		1	11	14	16	17	14	14	9	9	7	3	24
1994		1	12	15	17	20	19	15	13	14	13	8	29
1995		1	10	15	15	18	17	18	19	16	12	10	27
1996		10	16	16	16	16	17	18	21	15	11	8	36
1997		10	16	17	17	16	19	20	21	17	15	12	36
1998		9	17	16	16	18	17	15	10	12	11	8	27
1999	2	8	14	15	14	15	14	16	19	18	17		29
2000	3	13	17	18	19	18	17	15	14	14			30
2001		16	20	17	17	16	17	17	17				27
2002		15	16	16	16	17	19	20					27
2003	2	17	18	17	17	17	17						29
2004	1	15	17	16	15	16							21
2005	2	15	15	18	18								19
2006	1	9	9										9
2007	2	7											7
2008	1												1
Total	14	147	223	273	294	289	271	249	215	174	126	86	530

Note: N. Player gives the total number of players that belong to a cohort. In total, there are 530 players and 2,361 observations for which players are not in Young PSV and First team PSV. Players can only belong to one cohort, but can stay within the youth academy for multiple years.

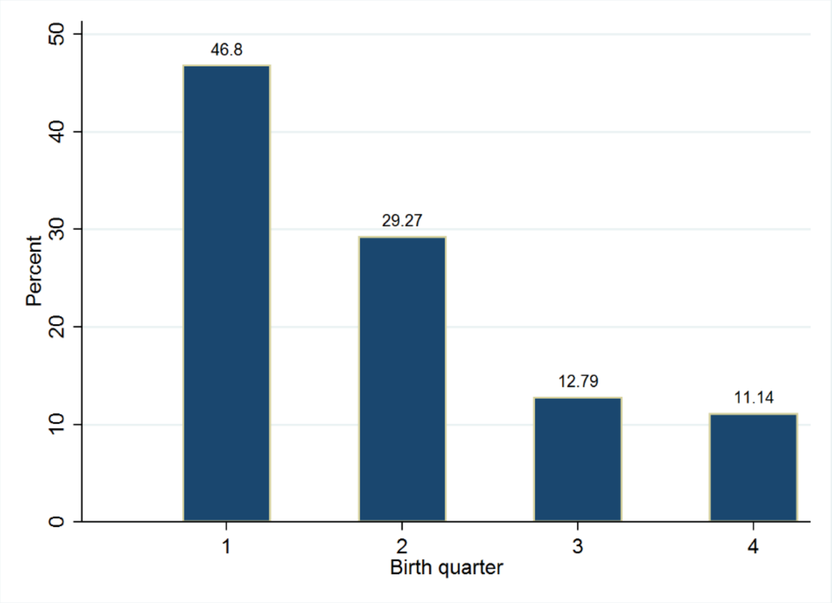


Figure 4.1A: Birth-date distribution by quarter

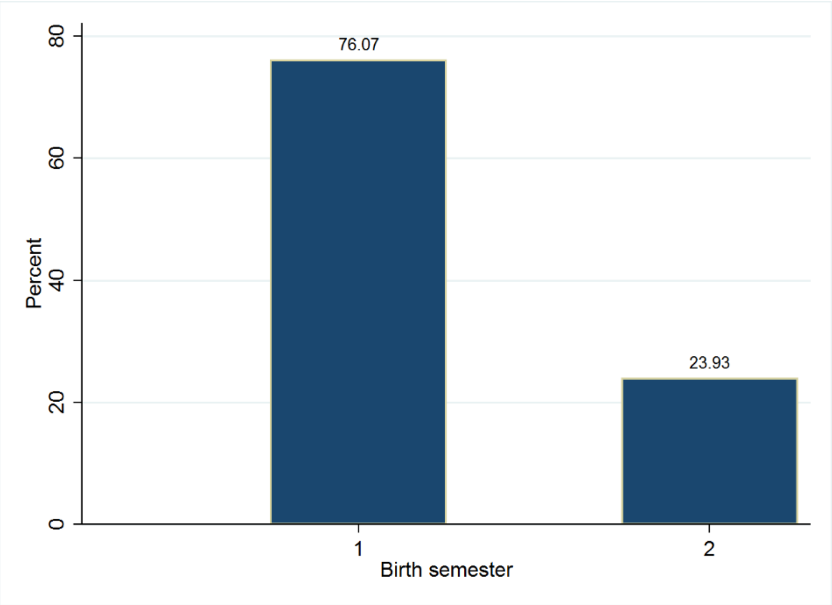


Figure 4.1B: Birth-date distribution by semester

4.5 The RAE

The RAE often is related to the biased selection of players at early ages. The recruitment of new players into the academy is necessary for the youngest age group, since a whole new team of players needs to be found. For the other youth teams, external selection is less urgent, but takes place in order to improve playing talent within the academy. Then, if playing talent is uniformly distributed across birth dates, one would expect a rather uniform distribution of birth dates within the academy as well.

The birth-date distributions are presented in Figures 4.1A, for birth quarters, and 4.1B, for birth semesters. The figures are comprised of the observations in Table 4.3. For both figures, a clear overrepresentation of early-born players is visible, i.e. players born relatively close to the cut-off date of January 1 are overrepresented. Although it follows from Figure 4.1A that the difference between the third and fourth quarter is rather small, we generally observe a declining pattern throughout the year. This typically points at an RAE that results from a selection system in favour of early-born players.

Table 4.4: Birth-date distribution by quarter

Quarter/Age	8	9	10	11	12	13	14	15	16	17	18	19	Total
Q1	5	65	113	126	141	137	127	117	101	80	56	37	1,105
Q2	6	50	61	81	81	85	77	67	65	55	38	25	691
Q3	3	22	30	33	32	33	33	34	27	24	18	13	302
Q4	0	10	19	33	40	34	34	31	22	15	14	11	263
Total	14	147	223	273	294	289	271	249	215	174	126	86	2,361
χ^2	6	51.9	95	88	101	102	88	77	76	61	36	20	788
p-value	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

To find out whether the RAE is persistent across ages, Table 4.4 gives the distribution of birth quarters for the different age groups. For all of them, an overrepresentation of early-born players is visible. With a χ^2-Goodness-of-fit-test, I test whether the observed distribution is different from what may be expected in case selection is based on (perceived) talent and talent is uniformly distributed across birth-dates.[8] Take the age of 17 as an example. The observed distribution by quarter is 80, 55, 24 and 15, which sum up to 174. Under the assumption that talent is independent from birth-date and if the youth academy aims to attract and develop talent, we would expect an equal number of players from each birth quarter, i.e. 43.5.[9] Taking the differences between observed numbers and expectations, results in the distribution 36.5, 11.5, -19.5 and -28.5. This indicates that more players than expected are born in the first two quarters of the year, while less than expected are born in the final two quarters. The χ^2-test is able to reveal whether these differences are statistically significant. It follows from the last two rows of Table 4.3 that the differences are significant, except for age category 8, which suffers

[8] For simplicity, I assumed that actual birth-rates are uniformly distributed (or more specifically, supply of players is uniformly distributed across birth-dates) which is common in this type of studies where multiple years are grouped together. As a check, I searched for actual birth-rates at Statistics Netherlands. They provide data on bird-rates per quarter as of the year 1995. Taking into account all births in the period 1995-2008, a total of 2,716,876 (including both boys and girls), I find the following birth-rates per quarter: 0.244 (Q1), 0.247 (Q2), 0.263 (Q3) and 0.246 (Q4). For semesters the rates are: 0.491 (H1) and 0.509 (H2). It follows that the differences are small, with a slightly higher share of births in the second part of the year. If any, this suggests that, under the assumption that talent is exogenous to birth dates and selection is based on talent, we would expect a slight overrepresentation of late born players. Thus, I think it is safe to assume a uniform birth-date distribution in my analysis.

[9] Naturally, a half person is impossible, but this mathematical result has no consequence for the statistical inferences.

from a low number of observations. In general, Table 4.4 reveals that the RAE is statistically significant and persistent across ages.

Table 4.5: Birth-date distribution by semester

Semester/Age	8	9	10	11	12	13	14	15	16	17	18	19	Total
H1	11	115	174	207	222	222	204	184	166	135	94	62	1,796
H2	3	32	49	66	72	67	67	65	49	39	32	24	565
Total	14	147	223	273	294	289	271	249	215	174	126	86	2,361
χ^2	4.6	46.9	70	73	77	83	69	57	64	53	31	17	642
p-value	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ratio	3.7	3.6	3.6	3.1	3.1	3.3	3.0	2.8	3.4	3.5	2.9	2.6	3.2

A similar conclusion can be drawn from the results in Table 4.5, where semesters are used instead of quarters. Again, the differences between the observed distributions and the expected number of players are significant, as follows from the results of the χ^2-test statistic. Furthermore, the last row of Table 4.5 gives the ratio of the number of players born in the first semester and the number of players born in the second semester (i.e. H1/H2). This ratio serves as an indicator of the overrepresentation of early-born players. A value of one indicates an equal number of players. However, the ratios are all close to three, which suggest that the academy contains three times more early-born players than late-born peers. Figure 4.2 shows the development of the ratio for the different age groups. Rather high values are followed by a gradual decline. At the age of 16, the ratio increases again and remains rather high for the age of 17. Finally, it reaches the value of 2.6 for the players aged 19. In general, the pattern follows a downward trend. This points at a reduction of the overrepresentation of early-born players and is in line with the deprecation of the physical advantage of these relative older peers. However, there is also a steep increase visible at the age of 16. This spike is at odds with the reduction of relative age differences over time, and therefore, does not fit within the framework of the RAE. It suggests an increased preference for early born players at the ages of 16 and 17 compared to both the age of 15 and 18. In search for an explanation, I find in Table 4.3 that the fluctuation in the number of players per age is rather high for the ages 16/17. This, at least partly, results from the changes in the number of teams within the academy over time (i.e. B1 and B2 vs. B1 only). Different selection procedures might arise with only one team compared to two. A further investigation reveals that at the age of 16, the fraction of early born players for the age cohorts 1988/1994 is 0.74 and does not differ much from the fraction for the other cohorts (0.79). For the age of 17 these fractions are 0.77 and 0.78 respectively. Furthermore, they do not differ much from those for the age of 15 (0.70 and 0.77) and the age of 18 (0.74 and 0.76). The difference in selection procedures, thus, does not seem to drive the spike and since I cannot think of an alternative explanation that fits with the data, the increased preference for early born players remains unexplained in this study and something that might be clarified in discussion with staff members of the youth academy. In such discussions, one should note that the vertical axis in Figure 4.2 ranges between 2.5 and 4. The jump in the graph from age 15 (ratio of 2.8) to the age 16 (ratio of 3.4) is 0.6. This is an increase of about 21 percent. This might be quite substantial, but not as huge as the spike may suggest. In general, it should be noted that the overrepresentation of early born players, measured by the ratio H1/H2, remains well above unity. Thus, although the RAE decreases, it is persistent.

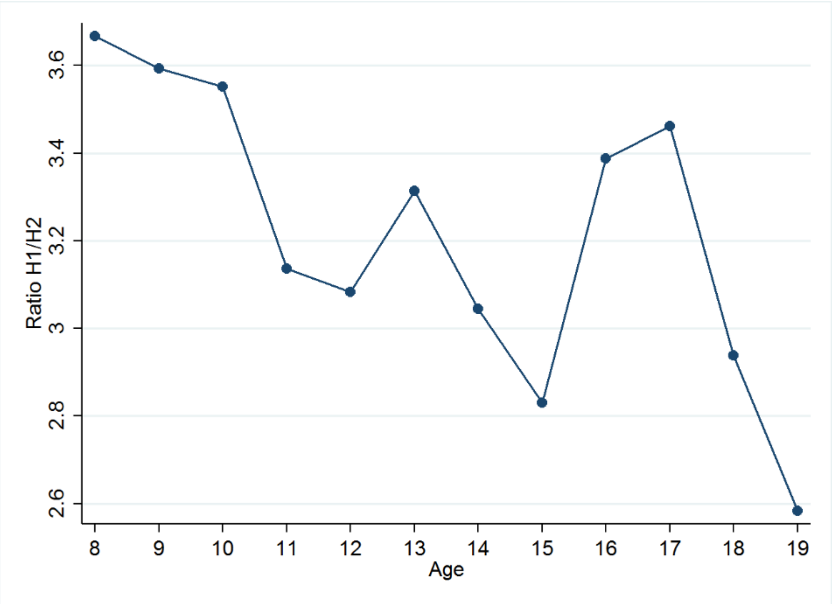


Figure 4.2: Ratio of players born in the first and second semester by age categories

4.6 External and internal selection

In the previous section, I showed that the RAE is persistent within the youth academy of PSV. Obviously, this results from the selection system that is applied. In general, and as discussed in the introduction, we can distinguish between two types of selection mechanisms. First, external selection deals with the recruitment of players from outside. Second, internal selection focusses on the players who are already active in the academy. In each season, the club can decide whether a player may stay or has to leave the club. In this section, I investigate how both types of selection relate to the occurrence and persistence of the RAE.

Assume that talent is uniformly distributed. Then, an overrepresentation of early-born players suggest that some relatively untalented age-advantaged players are selected in favour of some more talented age-disadvantaged peers. This may result from external selection, where talent has to be inferred, based on a rather low number of observations. In contrast, for internal selection, daily observations should result in a rather accurate idea of a player's talent, since the development of skills and performances can be monitored more easily. In that sense, it may be expected that the overrepresentation of early-born players, as results from external selection, is reduced by the internal selection mechanism.

Table 4.6: Overview of players who enter and leave the academy by birth semester

New-Out/Age	8	9	10	11	12	13	14	15	16	17	18	19	Total
New H1	11	105	72	55	30	35	21	22	21	13	7	4	396
New H2	3	29	21	24	12	8	10	11	5	3	5	3	134
Ratio	3.7	3.6	3.4	2.3	2.5	4.4	2.1	2.0	4.2	4.3	1.4	1.3	3.0
Out H1	1	13	22	15	35	39	42	39	43	42	23		314
Out H2	0	4	7	6	13	10	13	21	13	10	7		104
Ratio		3.3	3.1	2.5	2.7	3.9	3.2	1.9	3.3	4.2	3.3		3.0

Note: New players enter the academy at the start of the season in the given age category. The reported numbers of players that leave the academy, represent the number of players that are forced to leave at the end of the season for the given age category. Players who are promoted to Young PSV or to the First team of PSV are not included. Since all players have to leave the academy after the age of 19, we do not report numbers for this group. It follows that 530 players enter the academy. 418 have to leave at some point in time. For the remaining 112 players, 28 are promoted to one of the senior teams of PSV, while 84 leave the academy after the age of 19, when the academy ends. They either go to a senior team of PSV or move somewhere else.

First, I investigate whether the external selection mechanism indeed results in a biased birth-date distribution. The first part of Table 4.6 shows the number of players that enter the academy for different age groups, split by birth semester. It follows that, for all ages, the majority of players that enter the academy is born in the first half of the year, i.e. New H1 is larger than New H2 for all ages. This is also reflected in the value for the ratio, defined as New H1/New H2. A value of one indicates that the number of entrants that is born in the first half of the year is equal to the number of players that is born in the second half of the year. Only at the ages of 18 and 19, the value becomes close to one. However, for the ages of 16 and 17, the ratio is well above four. This, thus, means that the relatively older players are more than four times overrepresented in the group of new players. In general, the findings suggest that the external selection system indeed favours the age-advantaged players over their age-disadvantaged peers.

The second part of Table 4.6 shows the number of players that leave the academy as a result of internal selection. Note the players leave at the end of the season for a given age group. Furthermore, since all players have to leave the academy after the age of 19, I do not report any numbers for this age. For all the other age groups, the number for Out H1 is larger than Out H2. This means that the majority of players that leaves the academy is born in the first semester. The values for the ratio Out H1/Out H2 reveal that the number of early-born players who exit the academy is about three times as high as the number of late-born players. Interestingly, the overall ratio for entrants is three as well. This suggest that the share of early-born players that enters the academy is about equal to the share of early-born players who leaves the academy. Figure 4.3 shows that this seems to be the case throughout ages.

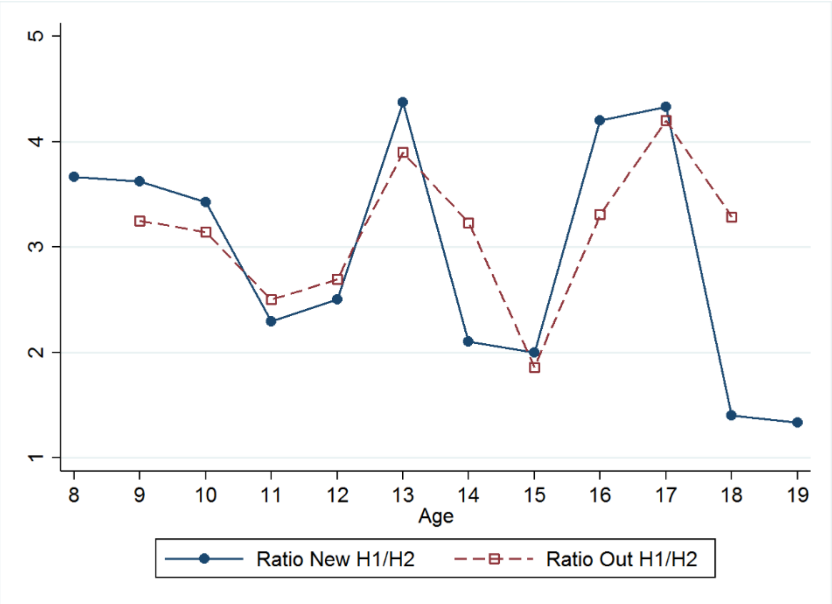


Figure 4.3: Ratios of players born in the first and second semester for new players and players who leave the academy

Based on the previous findings, I conclude that the internal selection system does not contribute to the overrepresentation of the age-advantaged players. However, it does not result in a (severe) reduction of the biased birth-date distribution either. The equal proportions of early-born players that enter and leave the academy, suggest that the selection bias that results from external selection is not overcome by the internal selection mechanism. To further investigate the influence of internal selection, I look at the probability of being in the academy for one additional year, two additional years and three additional years after the year that someone started. In particular, I investigate whether the probabilities are different for players who are born in the first semester and players who are born in the second semester. Since I want to cover the full range of ages, I restrict the sample to the cohorts of players who are born in or prior to the year 1998. Furthermore, I use restrictions on the starting ages of players, so that the final age of 19, at which the academy ends, does not serve as a boundary. Goalkeepers are excluded, because they do not compete with the other (field) players for a position in next season's squad. In general, each youth squad contains two goalkeepers. Finally, a few observations are excluded, since these players were promoted to one of the senior teams of PSV in either the first, second or third year after they entered the academy. Table 4.7 provides some descriptive statistics. The variable *+1 year* represents a binary variable that indicates whether a player remains at least one additional year in the academy after his first year. Similarly, the variables *+2 years* and *+3 years* are binary variables with the value of one if the player at least stays two additional years or even three years, respectively. It follows that 91 percent of the players stays at least one additional year, while this value drops to 70 percent for three additional years. Furthermore, the table shows the mean values for three other variables, split by the players who continue in the academy and those that leave the academy. Of main interest are the values for H1, a dummy variable that indicates that the player is born in the first half of the year. The value 0.75 for *+1 year* means that 75 percent of the players who remain for at least one additional year, is born in the first semester. In comparison, 88 percent of the players who leave the academy is early-born. The difference between the two values suggests that internal selection favours the late-born players. Although the difference is reduced for the other two variables that measure the number of additional years, it is rather stable. In a similar way, the differences for Dutch, a dummy with value one if a player has the Dutch nationality, suggest that Dutch players remain longer in the academy, compared to their non-Dutch peers. Finally, the results for starting age suggest that someone, generally, stays longer if he starts at an early age.

Table 4.7: Mean values for additional years in the academy

	+1 year		+2 years		+3 years	
Observations	271		265		253	
Starting Age	≤18		≤17		≤16	
Mean	0.91	0.09	0.80	0.20	0.70	0.30
	Yes (N=246)	No (N=25)	Yes (N=213)	No (N=52)	Yes (N=177)	No (N=76)
Dutch	0.70	0.64	0.71	0.65	0.74	0.62
Starting Age	11.93	13.84	11.79	12.71	11.55	12.18
H1	0.75	0.88	0.75	0.83	0.74	0.80

Note: Samples include players who are born in or before 1998 and do not include goalkeepers and players who are promoted to senior teams during junior ages. Furthermore, +1 year, +2 years and +3 years are dummies with value one if someone is in the academy at least one additional year, at least two additional years or at least three additional years after the first year, respectively.

To test whether the differences are statistically significant, I estimate a linear probability model of the form:

$$Y_i = \eta_c + \lambda X_i + \varepsilon_{ic} \qquad (1)$$

in which Y_i is the dependent variable. It is a binary variable for either one, two, or three additional year(s) in the academy for individual i. Furthermore, X_i is a vector of variables that includes dummies for a player being Dutch and whether the player is born in the first semester. In one of the model specifications, it also includes dummy variables for the starting age of the individual player. The vector of coefficients is given by λ, while ε_{ic} is the error term. Finally, η_c represents fixed cohort effects that control for unobserved heterogeneity between the cohorts, such as quality differences. As an alternative, I also use cohort * starting age fixed effects (i.e. an interaction between cohort effects and starting age effects). With this type of fixed effects, I explain the variation in the dependent variable by variation between peers from the same cohort, who started at the same age in the academy. For these players, the club has had an equal amount of observations and time to track the development of skills and to evaluate whether someone may stay within the academy or must leave.

Table 4.8: Parameter estimates internal selection

	(1) +1 year	(2) +2 years	(3) +3 years	(4) +1 year	(5) +2 years	(6) +3 years	(7) +1 year	(8) +2 years	(9) +3 years
Dutch	0.035	0.063	0.127	0.007	0.059	0.140*	0.023	0.095	0.082
	(0.037)	(0.067)	(0.077)	(0.037)	(0.059)	(0.065)	(0.050)	(0.067)	(0.074)
Starting Age 10				-0.103*	-0.024	0.013			
				(0.051)	(0.062)	(0.041)			
Starting Age 11				-0.066	0.009	0.074			
				(0.039)	(0.082)	(0.068)			
Starting Age 12				-0.055	0.103	0.141			
				(0.051)	(0.070)	(0.102)			
Starting Age 13				-0.091	0.093	0.152			
				(0.057)	(0.072)	(0.085)			
Starting Age 14				-0.107*	-0.156	-0.008			
				(0.054)	(0.110)	(0.114)			
Starting Age 15				-0.194**	-0.052	-0.287***			
				(0.078)	(0.131)	(0.078)			
Starting Age 16				-0.114	-0.049	-0.018			
				(0.090)	(0.089)	(0.098)			
Starting Age 17				-0.543***	-0.397**				
				(0.130)	(0.151)				
Starting Age 18				-0.200					
				(0.153)					
H1	-0.067***	-0.071*	-0.089	-0.064**	-0.071*	-0.100	-0.075*	-0.118**	-0.171**
	(0.014)	(0.038)	(0.066)	(0.023)	(0.039)	(0.075)	(0.038)	(0.058)	(0.075)
Constant	0.935***	0.814***	0.677***	1.056***	0.834***	0.650***	0.949***	0.828***	0.772***
	(0.022)	(0.056)	(0.079)	(0.037)	(0.068)	(0.060)	(0.038)	(0.067)	(0.080)
Observations	271	265	253	271	265	253	271	265	253
Starting Age	≤18	≤17	≤16	≤18	≤17	≤16	≤18	≤17	≤16
Cohort effect	11	11	11	11	11	11			
Cohort*Starting Age effect							84	80	73

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.* All estimates are based on player that are born in 1998, not including goalkeepers and players who are promoted instead of left. Dependent variable is a dummy that indicates whether a player is still in the academy +1, +2 or +3 years after his first year. Models (1)-(6) contain standard errors clustered by cohort, models (7)-(9) contain standard errors clustered by cohort*starting age.

The results are presented in Table 4.8. In the models (1)-(3), I control for Dutch players (coefficient is insignificant in almost all specifications) as well as cohort effects. The result for H1 is negative and highly significant in model (1). Furthermore, in model (2) the coefficient is negative and significant only at the 10%-level, whereas the coefficient is negative and insignificant in model (3). Adding dummies for the starting age of a player in specifications (4)-(6) does not change the results for the coefficients of H1. However, some differences arise if cohort * Starting Age interaction effects are used instead of cohort fixed effects. Then, the coefficient for H1 in the model with *+3 years* as dependent variable becomes significant as well. Furthermore, the estimated values are all more negative. In general, these results suggest that players born in the first half of the year, have a higher probability of about seven percent to leave the academy in the first year after the year of entering, compared to their late-born peers. For a two-year time span, the probability is about the same, but higher if the cohort * starting age interaction effects are used. Then, I also find a significant probability for the three-year period. The result means that the players who are born in the first half of the year, have a higher probability of about 17 percent to leave the academy within three years after their start, compared to players from the same cohort and starting age, that are born in the second semester.

Although the estimated values of the coefficients for H1 are not very high, it has some meaning. Overall, is reveals that the internal selection system works as predicted, i.e. reduces the overrepresentation of early-born players. However, the effect is modest and not strong enough to overcome the biased birth-date distribution. The external selection mechanism attracts too many age-advantaged players to make a difference in that respect.

4.7 Professional football

As a final part of analysis, I look at the players' career at the senior age level. More specifically, I focus on the distinction between players who have become a professional football player and those that have not been a professional. For most players, it is quite easy to determine whether he achieved a career as a professional. Some of them sign contracts at European top teams, such as Memphis Depay, who left PSV for the English Premier League giant Manchester United FC and later moved to Olympique Lyonnais in the French Ligue1. Other players end up at the lower levels of Dutch amateur football. Clearly, the former group contains professionals, the latter contains the non-professional players. However, in some cases it is more difficult to determine. For example, when someone is part of the squad of a professional football team, but only plays a small number of matches. I tried to tackle this issue and construct a measure for professional football that is based on information obtained from various internet sources such as transfermarkt.de, vi.nl and many club websites. Furthermore, newspaper archives are used. In general, I decide to consider a player as a professional in case he belonged to the squad of a club that played on the first or second highest level in their domestic league. Since Young PSV is playing at the second level in the Netherlands as of the 2013/14 season (before that season, Young teams were not allowed to participate in the professional leagues) many players who belong to Young PSV obtain some experience in professional football. However, within this study, I do not consider this as professional football. First, because the players who belong to Young PSV are, in some sense, extending the academy in anticipation of promotion to the first team. Furthermore, Young PSV only plays on the second level as of the 2013/14 season, which is a second reason not to consider this as professional football. The players who are born in the earlier seasons of the sample and played for Young PSV, did not have the opportunity to play on the professional level, because this was not possible yet.

Since players may enter professional (senior) football at different ages, I distinguish between the ages 19, 20, 21, 22 and 23. For each age, I construct a binary variable that indicates whether the person was a professional player. In doing so, I take account of those players who play professional football early in their career, but leave the professional level after only a few years, prior to the age of 24. The variables are used in an analysis that is quite similar to the one applied in the previous section. Table 4.9 provides mean values for the five different ages. Note that for each age, a restriction is used on the age cohorts that are included. The restrictions assure that the samples only contain players who at least reached the age as given in the dependent variable during the 2016/17 season. Furthermore, goalkeepers are excluded again. The first columns show that about 1/6 of the players were a professional at the age of nineteen. For the age of twenty, this is approximately 1/4, while about 1/3 for the other ages. Furthermore, in the last row, the share of players that is born in the first semester is given. It follows that, for all ages, this share of early-born players is above 0.5. Unsurprisingly, and in line with the overrepresentation of the age-advantaged players in the academy, most professionals are born in the first semester. This is, however, also true for the ones who did not become a professional player. For the ages 19, 20 and 21, the share of early-born players is lower for the professionals than the non-professionals. This suggest that the early-born players do relatively worse, in terms of achieving professional football, compared to their late-born peers. The opposite is true for the ages 22 and 23. For these samples, the share of early-born players is higher for the professionals than for the non-professionals.

Table 4.9: Mean values for players who have become a professional

	Professional age 19		Professional age 20		Professional age 21		Professional age 22		Professional age 23	
Observations	284		260		227		197		173	
Birth year	≤1998		≤1997		≤1996		≤1995		≤1994	
Mean	0.15	0.85	0.24	0.76	0.33	0.67	0.35	0.65	0.31	0.69
	Yes (N=42)	No (N=242)	Yes (N=62)	No (N=198)	Yes (N=76)	No (N=151)	Yes (N=68)	No (N=129)	Yes (N=54)	No (N=119)
Dutch	0.57	0.71	0.66	0.69	0.64	0.70	0.66	0.66	0.63	0.67
H1	0.62	0.78	0.68	0.76	0.70	0.76	0.75	0.71	0.78	0.70

Note: Samples do not include goalkeepers. Professional age 19, 20, 21, 22 and 23 are dummies with value one if someone is a professional football player at that age.

To test whether the differences are statistically significant, I use a similar linear probability model as given by Equation (1). As the dependent variable, I use the five variables for the different ages. These have the value of one in case a person was a professional player at that age. Furthermore, I include a dummy for Dutch players and cohort fixed effects. The variable of interest is the dummy H1 that indicates whether the player was born in the first half of the year. Table 4.10 shows the results. The dummy for Dutch is insignificant in all specifications. For H1, the coefficient is significant in specification (1) that models the probability of being a professional football player at the age of 19. The negative coefficient means that early-born players, compared to their late-born peers, have a lower probability of about 11 percent to be a professional at this age. Although the coefficients in models (2) and (3) are negative as well, they are insignificant. For the ages 22 and 23, the coefficient becomes positive, but remains insignificant. Thus, only at the age of 19, when players actually still are youth players, I find a statistically significant difference. The difference is in favour of the late-born players and points at a reversal of the RAE. A possible explanation may be that, at early ages, those late-born players learn how to deal with more developed older aged peers. This may be beneficial later on in their career and is an experience that early-born players, generally, do not obtain. Furthermore, it may be that the late-born players within the academy, on average, are highly talented, compared to less talented players who are in the academy. Otherwise, they would not have managed to survive up to the senior teams. In that respect, the system only selects the high potentials from the group of late-born players. At the same time, the early-born players are overrepresented in the group of professionals. Combined, this suggests that it would be beneficial to select some more modest talents from the group of late-born players, who are currently not selected. These then can replace the worst early-born players. The result will be a more uniform birth-date distribution.

Table 4.10: Parameter estimates professional football

	(1) Professional age 19	(2) Professional age 20	(3) Professional age 21	(4) Professional age 22	(5) Professional age 23
Dutch	-0.110	-0.044	-0.056	-0.007	-0.033
	(0.067)	(0.079)	(0.067)	(0.062)	(0.053)
H1	-0.108**	-0.085	-0.056	0.047	0.084
	(0.035)	(0.058)	(0.061)	(0.059)	(0.064)
Constant	0.305***	0.332***	0.414***	0.316***	0.273***
	(0.050)	(0.063)	(0.063)	(0.056)	(0.069)
Observations	284	260	227	197	173
Birth year	≤1998	≤1997	≤1996	≤1995	≤1994
Cohort effects	11	10	9	8	7

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1*. Estimates are based on player observations that are born in or before the year 1998 (1), 1997 (2), 1996 (3), 1995 (4) and 1994 (5), not including goalkeepers. Dependent variable is a dummy that indicates whether a player is a professional at the age of 19, 20, 21, 22 or 23. All estimates contain cohort effects.

4.8 Discussion

This section will provide a discussion. First, I will deal with some assumptions that were made, in particular on the supply of potential talent. Second, the generalizability of results is discussed. Here, I address the sport-dependence of the selection system, with a special interest in women's football. Furthermore, I discuss how my results may apply in the setting of an educational system.

4.8.1 Assumptions

The introduction started with a simplified characterization of the labour market for top-level talents, with low demand for high potentials and high supply of people. Although this in general could be true and applicable to the situation of youth professional football, I made some more explicit assumptions in my analyses. First, I assumed that the supply of players is uniformly distributed across birth dates (see also footnote 8). We saw that birth-rates are about uniformly distributed, but it might be that supply is not. For example, because late born and disadvantaged players do not play football in equal portions as their early born peers. Potentially because they are not able to obtain (sufficient) access to training facilities as a result of selection that takes place at amateur teams. Related, they might also lack the ambition to reach their potential level of performance. I cannot rule out that this is happening and lack data (i.e. birth-date distributions) on the supply side of the market for players. However, football is very popular within the Netherlands, with a high density of football clubs across the country, especially compared to other sports clubs. Many young boys start to play football at some age and all have comparable access to facilities. Then, if supply consist of all players who are active at a football club, there will be plenty of players available for the professional teams, with birth dates throughout the year. And, as far as this birth-date distribution is skewed in favour of early born players, which means the supply of players for PSV is skewed, we can say that the selections system of the academy is not able to correct for the relative age difference that generally exists within youth football.

We saw that some correction of the RAE takes place via the internal selection system. I made an implicit assumption about this internal selection system. I assumed that PSV selects such that they are left with the most talented players. Thus, each decision is made carefully, given the available options. A careful decision means that each player is evaluated on a yearly basis and the decision to keep someone within the academy is based on the idea that this person might be able to reach until the first team. However, both coaches and players know that not all players will reach that level. Primarily, because only few are needed each year. And if it is clear at young ages who those will be, it might not be necessary to be strict about the selection process, since then it does not matter who is playing together with these high potentials. There would be no need to find an outside replacement for someone who already plays within the academy. This could be easier, and I cannot rule out that this is happening. However, since it will be difficult to evaluate who will be the high potentials and since this may change over time, especially at young ages, it seems advisable to remain strict. Therefore, I suggest my assumption seems plausible, at least on average.

The final assumption to be discussed, relates to constraints that PSV faces within their selection procedures. I assumed that the decision to play for PSV or leave the academy was always made by the club. However, players may prefer to play for other teams and might also leave for other reasons. For example, because of an injury, upward mobility, or simply because someone is not willing to play for PSV anymore. I have some information about upward mobility (e.g. Patrick van Aanholt who moves to the youth academy of Chelsea FC) but it seems likely that information on (voluntary) quits is far from complete. Especially for players who left PSV at early ages and did not become famous. It seems quite difficult to obtain such information. My guess would be that, within the group of players who left the academy, more did so on their own initiative and not because they were forced by PSV. In that respect I have incomplete data on those who left. My assumption that the decision to leave the academy was always made by PSV, therefore seems to be the second-best option. As far as the own-initiative-quits are random and unrelated to their birth dates, I suggest that this is not too much of a problem for my results, as far as the number of these quits is rather low.

4.8.2 Generalizability

The setup of this study is quite narrow in the sense that data is used of only one youth academy in male professional football. An advantage of this narrow scope is that it allows for an in-depth analysis of the selection process, including a distinction between external selection and internal selection. Although the results are obtained through a case study, i.e. the youth academy of PSV, it provides some interesting insights for youth football in general and future research in that field, covering other teams and leagues.

Besides that, I suggest that my findings are also of interest for other selection systems within sports. The RAE results from selection systems that do not correctly incorporate relative age differences. Therefore, my suggestion would be to select at an age, when relative differences, mainly related to physical aspects, are small. For gymnastics, in which late maturation is an advantage, this would be early in life (taking into account ethical issues), while for football around the age of 15. In sports where differences in maturation are less likely to influence performances, it does not seem to matter much (e.g. shooting).

Furthermore, the results for the internal selection process revealed that an increased number of observations of individual performances, compared to few external observations, seems to result in a better prediction of potential talent of young players. This means that, at young ages, one should not focus too much on a small group of individuals. Instead, one preferably monitors the development of a rather large group of athletes, before selection takes place. For this to work, two other aspects seem relevant. First, the amount of competition for talent and, thus, the need to outperform your competitors. For example, football clubs that compete for a promising player. Second, and related, the organization that performs the selection. Selection may primarily happen at (private) professional clubs as is the case in football. Selection may also happen at the sports body, as with ice skating and track and field. Differences between these types of organizations are the availability of funds and the availability and spread of training facilities within the country. Private clubs generally have more funds available to select earlier in life and have to compete against each other. Sports bodies have less funds to finance their talent selection, identification and development programs. But, they have a monopoly and may help local clubs (i.e. their members) with the provision of facilities. In that way, a rather large group of athletes may receive similar support until an older age at which selection takes place.

The selection process may also differ in the relevance of certain components for certain sports. Within football, one has to rely on subjective evaluations of scouts and coaches. Especially with external selection, this may result in a rough approximation of talent. With internal selection, more information is available and might be gathered objectively, e.g. physical development and even mental status. Development in other aspects, such as tactical and technical skills, can only be subjectively evaluated within the context of a team effort. This differs for sports such as ice skating and track and field, in which each individual obtains an objective measure of performance. This can be a time, a distance or a height. These performances are used in the selection process and make it easy to find the best performers. Individuals that are not yet selected for additional training and support can qualify quite easily, since a clear comparison between performances of peers is available. However, as with football, too much focus on these performances may result in an overrepresentation of early born players. For all sports, coaches should incorporate relative age differences and adjust performances based on birth dates. This might be easier with objective performance measures than with subjective evaluations. In both cases, there remains a role for the coach within the selection process. And if they have sufficient observations of individuals, as with internal selections, I suggest they do rather well in their evaluation of potential talent.

It seems likely that the same reasoning applies to female sports. In the following, I focus on women's football, which is rapidly growing in the Netherlands. Most female players who are active in the highest tier of Dutch football are semi-professionals. A further difference with male football, and related to this study, is that the professional clubs do not have an academy for females. Thus, the selection of female players takes place at a relative old age, i.e. not at very young ages such as with boys. For this selection, they have to rely on amateur teams and, of course, the other professional clubs. Since selection by the professional teams takes place at an older age, in general, after maturation, the relative age differences might be less severe. However, it could of course be the case that the RAE exists at the amateur teams. Supply of players comes from these amateur teams. As far as the number of these teams is rather high, the talent pool remains large until adolescence (i.e. ages 16-19), while all players may receive equal opportunities to develop. This in contrast to male football, where the talent pool at the ages 16-19 mainly consists of the players within the youth academies of the professional clubs. For males, it is unlikely to obtain a professional career if you are not playing in one of the youth academies at these ages. For females, no academies exist and, thus, this is not an issue. In that respect and related to the RAE, my suggestion is that the selection process within women's football is doing better, since the talent pool remains rather large until adolescence. A drawback of this selection method might be that additional facilities, practice and support are missing, because there is no professional youth academy. Amateur teams may provide this support, and in some way serve as regional training centres that closely cooperate with the professional clubs.

As a test for the RAE, I looked at two female adult selections and the distribution of birth dates. First, the 2017/18 female team of PSV consists of 22 players. Eight are born in the first quarter of the year, 7 in the second quarter, 6 in the third quarter and only 1 player in the fourth quarter. In terms of birth semesters, the distribution is 15 in the first half of the year and 7 in the second half of the year. Thus, there is an overrepresentation of early born players, but less severe compared to the male academy. Second, I also looked at the female Dutch national football team that became European Champions in the summer of 2017. Here, out of 23 players, 3 are born in the first quarter, 6 in the second quarter, 4 in the third quarter and 10 in the fourth quarter. In terms of birth semesters, this is 9 in the first half of the year and 14 in the second half of the year. Thus, the distribution of birth dates contradicts with the RAE and the main results of this study. Without drawing to many conclusions based on these single observations, something seems to work well with respect to the RAE within in the selection system of women's football.

Although this paper is focused on the RAE in elite youth football, I suggest that some results are interesting outside of this domain as well. In particular, I think that the distinction between external and internal selection systems, shows resemblance to the way in which children on primary school are evaluated. In the Netherlands, the children in the final grade all have to take an external test. The result that is based on this single observation prescribes the potential abilities and what level of secondary school should fit for the child. Secondary schools may base their admissions on these test scores. Besides this external test, teachers, as coaches in football, continuously evaluate the development of skills and abilities, which serves as an indicator of potential as well. Since relative age differences matter in school systems, I suggest that these internal evaluations of teachers should not be underestimated. Perhaps it would be good to integrate the external test score in the advice to the secondary school. Then, this school cannot differentiate between external and internal evaluations of potential and one can be sure that both are used. How this should work in practice, is something for future research. One could also think of the optimal age to move from primary school to secondary school. Currently this happens at about the age of 12. Based on the discussion above, I suggest to take maturation into account within this selection system and perhaps delay selection a little.

4.9 Conclusion

Previous studies find that selection systems for talent suffer from a bias in the selection of relative age-advantaged players. As a result, this group of peers is overrepresented, which is called an RAE. Under the assumption that talent is uniformly distributed across birth dates, the RAE suggests a loss of talent. In this paper, I investigate how this works for elite youth football, with data from the youth academy of PSV. The dataset contains all players who are born in 1988 or later and were active in the academy for at least one season. A birth-date distribution that combines all birth cohorts and ages, reveals that the academy consisted of about three times as many players who are born in the first half of the year, than players who are born in the second half of the year. Separate distributions by age reveal that the overrepresentation of these early-born players is persistent across ages. Despite that the phenomenon of an RAE is known for a few decades now, this result suggests that, in general, selection systems of top-level talent are still not able to correctly incorporate relative age differences. However, the analysis on the internal selection mechanism suggest that this functions rather well in that respect. Compared to external selection, the number of observations prior to the selection decision is high for internal selection. This makes the evaluation of talent more accurate. As expected, the results show that internal selection reduces the severity of the RAE. However, the reduction is not large enough to overcome the bias in birth-date distribution that results from external selection. Thus, an overrepresentation of early-born players remains within the academy. Therefore, it is not surprising that the majority of professional players is born in the first semester of the year. Interestingly, however, at the age of 19, late-born players have a higher probability to become a professional player. This suggests that the group of late-born players, on average, contains highly talented peers. Thus, an improvement could be made by selecting some additional late-born players instead of some of the early-born players. Furthermore, as a remedy to the RAE, I suggest that external talent selection should incorporate more tracing of development prior to the decision of recruitment. As is the case with internal selection. This may be difficult for individual clubs, since they have to compete with other professional teams for potential talented players and, thus, need to decide quickly. A solution could be to install several regional training centres, in which multiple professional clubs participate. This system contains similarities with the current situation of Dutch women's football. Players can start in these regional centres at early ages. The development of skills and abilities is monitored closely, but recruitment by the participating teams is only allowed after several years. For example, it may not be allowed before the age of fifteen. Clearly, this idea is not very concrete yet, but at least includes the element of tracing development, before the decision to select.

Chapter 5

Team heterogeneity and performance

(Joint work with Jan van Ours and Martin van Tuijl)

5.1 Introduction

Teamwork is common in everyday life. Both in working life as well as during leisure activities such as sports, people cooperate in teams, while aiming for a common goal. This goal may take on many different shapes, such as the completion of a project, the publication of an academic paper, or a match victory. Naturally, the outcome depends on various factors. Some of these factors are exogenous, for example, a budget constraint or luck, which is often claimed to be important in sports. Other factors can be influenced by individual team members and teams as a whole. Here, one can think of individual effort levels and team cooperation, team coordination and team composition. The eventual outcome, i.e. the team performance, results from a combination of all relevant elements. If all exogenous factors are given and human input is relevant, team performance is determined by the individual inputs of team members and the way in which these are combined. In general, team performance improves, if individual inputs increase. Furthermore, improving the way in which these inputs are combined results in better performances. Of course, both aspects can work simultaneously and/or in an interactive way. Take, for example, a change in the heterogeneity of a football team, which is the topic of the present study. Suppose a player is replaced by another player, who is more experienced. This may have an impact on the individual performances of their teammates. For example, because this more experienced player knows how to handle in certain difficult situations. Furthermore, it may impact team performance, because the individual input of this experienced player is weighted differently than other inputs.

Scholars have been interested in teams, team members and performances for a few decades now. Studies relate to incentives, contracts and payments, such as, incentives and discrimination in pay (e.g. Winter, 2004), the pay-performance relationship (e.g. Torgler and Schmidt, 2007; Bryson, Buraimo and Simmons, 2011) contract duration (e.g. Buraimo, Frick, Hickfang and Simmons, 2015) expectations as reference points (e.g. Bartling, Brandes and Schunk, 2015) and career prospects (e.g. Miklós-Thal and Ullrich, 2016). Others looked at the role of experience within teams (e.g. Huckman, Staats and Upton, 2009) and productivity, age and aging of teams and team members (e.g. Van Ours and Stoeldraijer, 2011; Börsch-Supan and Weiss, 2016). Furthermore, the concepts of free-riding, peer effects and effort levels are related to teams and their performances (e.g. Kandel and Lazear, 1992; Hamilton, Nickerson and Owan, 2003; Gould and Winter, 2009). Of particular interest for the present study are those studies that investigate the impact of team heterogeneity on performance. This heterogeneity is measured in different ways, such as team heterogeneity in ability (e.g. Hamilton, Nickerson and Owan, 2003; Franck and Nüesch, 2010; Papps, Bryson and Gomez, 2011; Gerhards and Mutz, 2017; Muehlheusser, Scheemann and Sliwka, 2016), team heterogeneity in nationality (e.g. Lazear, 1999; Brandes, Franck and Theiler, 2009; Haas and Nüesch, 2012; Kahane, Longley and Simmons, 2013; Maderer, Holtbrügge and Schuster, 2014; Gerhards and Mutz, 2017) and team heterogeneity as a combination of factors (e.g. Beck and Meyer, 2012).

Most theoretical work on teams and performances is based on general economic concepts (e.g. Kandel and Lazear, 1992; Lazear, 1999; Winter, 2004). In line with Kahn (2000), who acknowledges that (professional) sports are useful to study labour market phenomena, the use of sports data is popular for empirical verification of the theoretical predictions. In the present study, we investigate the impact of team heterogeneity on performance. We use data from professional football. Of course, differences exist between teamwork within football teams and teamwork in daily work operations. For example, a professional football team receives a large amount of (media) attention. However, there is also plenty of similarity, such as the interactions between team members and the division of tasks. Therefore, we argue that empirical studies on the relationship between team heterogeneity and performance in sports are useful outside of the domain of sports as well.

The work of Lazear (1999) can be seen as a theoretical foundation for this relationship. His model is based on the concept of a 'global firm' that contains team members of different cultures or countries. Lazear describes three conditions that should be met for a global firm to operate in a superior way. First, information or skill-sets should be *disjoint* across team members. Second, these sets should be *relevant* for the other group. Third, *communication* should be possible, i.e. the costs of communication should not be prohibitive. Within the global firm, teams exhibit a diversity of cultures or nationalities with disjoint information and skill-sets that are relevant to one another. The costs of integration and communication should be lower than the benefits of working in an international and diverse setting. According to Lazear's model, this is regularly the case. Therefore, we derive the hypothesis that team heterogeneity concerning nationality is beneficial for performance. However, previous studies using sports data, only find limited support for this hypothesis. Brandes, Franck and Theiler (2009) use data from the German Bundesliga for the seasons 2001/02 – 2005/06 and find a negative effect of team diversity within defensive players on team performance, while no effect for other positions. Kahane, Longley and Simmons (2013) use data from the National Hockey League for the seasons 2001/02 – 2007/08, excluding the season 2004/05 due to a player lockout. They find that benefits accrue from diversity in nationality and that these benefits are the greatest if the group of foreign players is rather homogeneous, i.e. if they stem from the same country. Haas and Nüesch (2012) find a negative effect of diversity in nationality on team performance within German Bundesliga football (seasons 1999/00 – 2005/06). Maderer, Holtbrügge and Schuster (2014) come to a similar conclusion from data of the season 2008/09 of the highest level of professional football for the five largest European leagues (England, France, Germany, Italy and Spain). Gerhards and Mutz (2017) use data from the highest level of the twelve best European football leagues according to the UEFA Team Ranking in 2011/12. Their dataset covers the seasons 2011/12 – 2015/16. They investigate, amongst others, the influence of diversity in nationality on team performance using club-season observations. They find this relation to be non-linear and conclude that heterogeneity is beneficial up to a certain level.

Although Lazear (1999) bases his model on diversity in nationalities, it seems applicable to other types of heterogeneity as well. In particular, the elements *disjoint*, *relevant* and *communication* seem important for teams with different levels of age/experience as well as for teams with heterogeneity in abilities.[1] Think for example of the potential problems in communication between younger and older workers, or between co-workers with certain specific skill-sets. Although the relationship between age/experience and performance (productivity) has been studied numerous times (e.g. Huckman, Staats and Upton, 2009; Van Ours and Stoeldraijer, 2011; Börsch-Supan and Weiss, 2016), the number of empirical studies that focuses on the diversity in age/experience and performance is rather small, in particular when sports data is used. In their study on team diversity within the five largest European football competitions, Maderer, Holtbügge and Schuster (2014) include a diversity measure for age. They find a negative effect on team performance. Beck and Meyer (2012) use a heterogeneity measure in which variables for tenure, overall tenure, age, nationality, experience and success are combined into one value per club-match observation for German Bundesliga matches. Their sample covers the eleven seasons 1992/93 – 2002/03. They find a negative effect of team heterogeneity on team performance. However, because variables are grouped into one measure, it is unclear which separate element drives this result.

Measures of heterogeneity in ability are generally constructed from output related values. This can be objective measures, such as production rates in manufacturing or (individual) match details on, for

[1] Naturally, other theoretical frameworks may relate to these different types of heterogeneity, such as a human capital framework (e.g. Franck, Nüesch and Pieper, 2011). Since our aim is to empirically investigate the relationship between team heterogeneity and performance, we focus on the empirical findings of previous studies, without an in-depth discussion of these alternative theories.

example, passing, running and scoring in sports. Furthermore, in sports, one can rely on subjective expert evaluations of individual performances. In their study on team production in a garment plant, Hamilton, Nickerson and Owan (2003) investigate the effect of heterogeneity in ability (measured as an individual output rate) on team production and find a positive effect. They relate this positive effect to benefits that result from mutual learning as well as peer pressure and norm-setting by high ability peers (see also Kandel and Lazear (1992) for a discussion on how peer pressure and norm-setting might work). The process of mutual learning assumes that there is some degree of information sharing possible between agents, which increases substitutability between one another. Other studies indeed show that benefits of team heterogeneity arise in the situation where inputs (i.e. actions or tasks) are substitutes and not complements (e.g. Prat, 2002; Gould and Winter, 2009). Franck and Nüesch (2010) exploit this distinction between complements and substitutes in an analyses of team productivity in German Bundesliga football for the seasons 2001/02 – 2006/07. They proxy talent with objective individual performances and expert evaluations. Furthermore, they define a competition team, i.e. those players who participate in competition matches, which consists of complementary talents. For the sake of comparison, they define the entire squad, which also includes all players who do not participate in competition matches and where substitutability is larger. The authors argue, that the outcomes of competition matches are the result of complementary skills, while end-of-season outcomes result from the composition of the entire squad, in which learning takes place and team members are substitutes. Their hypotheses are that talent disparity within the competition team has a negative effect on team performance, while talent disparity within the entire squad has a positive effect on team performance. In general, their empirical analyses provide support for both. Papps, Bryson and Gomez (2011) find support for a non-linear relation between skill dispersion and team performance. They use a large dataset from the Major League Baseball covering the seasons 1920 - 2009. Skill is measured by conventional performance indicators for batting and pitching. They find that there is an optimum concerning heterogeneity for both batters and pitchers, implying that diversity in playing talent ought to be moderate. Gerhards and Mutz (2017) also investigate a non-linear relationship between the inequalities of playing talent (measured by market values retrieved from transfermarkt.de) and team performance. They use club-season observations and find insignificant results for both a linear specification as well as a quadratic form. The study by Muehlheusser, Scheemann and Sliwka (2016) is somewhat different and provides indirect evidence against the hypothesis of benefits of team heterogeneity. They investigate the differences in performance effects of coach-changes for homogeneous and heterogeneous teams. Notably, they measure heterogeneity by differences in expert evaluations. Their results suggest that benefits of coach-changes arise for homogeneous teams, in which competition between team members is large, but less so for heterogeneous teams, in which competition is rather modest.

From the above it follows that empirical evidence for the relationship of team heterogeneity and performance moves into different directions. This might have to do with the level of analysis (individual vs. team) as well as with the way in which heterogeneity and performances are measured. Performance measures, in particular in sports, might suffer from the fact that they are rather discrete. For individuals, this is the case for expert evaluations as well as crude performance indicators, such as goals scored and minutes played. On a team level, this holds for the number of points, goal difference, victory or defeat and (final) rankings. This is generally less of a problem in business, where there is, most likely, always some production rate or project performance. Within consultancy, for example, there is a continuum (of quality) of advice. However, these performances are often hard to relate to individual contributions. Furthermore, it might be difficult to construct proper heterogeneity measures, since bibliographical data and background information on individuals is missing (e.g. of past experience). This, in turn, is less of a problem for professional sports.

In our study we use such sports data, specifically from the highest tier of Dutch professional football. Our dataset covers the 2014/15 season. It includes detailed data on players and their playing histories as well as on matches and performance.[2] For individual performance we use data provided by ORTEC Sports. This is a company that collects detailed data on individual player actions during matches. Furthermore, these actions are classified in terms of successful/unsuccessful. We thus have a Success Ratio of actions for each individual player per match that serves as a performance indicator. As an alternative we also have the ratio of successful passes (Passing%). Using information retrieved from other sources, we construct four different types of heterogeneity measures. First, we construct a measure of heterogeneity in ability (skill) measured by market values. Second, we obtain a heterogeneity measure based on nationality. Third, based on the height of players, we also construct a heterogeneity measure of physical attributes of players. Finally, we use age and playing history to construct a heterogeneity measure of age/experience.

Our analysis consists of three parts. First, we start with the relationship between heterogeneity of team members and individual performances. We find a statistically significant effect for heterogeneity in nationality that suggests that diversity is beneficial. Also, we find a statistically significant effect for heterogeneity in experience, which suggests that diversity has a negative effect on performance. We show that this result depends on the use of club fixed effects. Furthermore, we use sensitivity

[2] A detailed description of our data is provided in the next section and in Appendix E.

analyses to investigate whether these results are stronger for certain subgroups and subsets of players. In general, we find robust results for all types of measures of heterogeneity.

In part two of our analyses, we relate the types of team diversity to team performance using club-match observations. For team performance, we use crude and discrete match-outcomes (i.e. the number of points, goals difference and winning) as well as continuous performance measures based on a team average Success Ratio and a team average Passing%. For team diversity in nationality and team diversity in experience, the results are similar to our findings for individual performance. Interestingly, the former is only significant for our more subtle measures of team performance that are based on average individual performances. Therefore, we conclude that team diversity can be related to team performance in a direct way, observable in match-outcomes, but might also indirectly relate to team performance via an improved average Success Ratio.

Based on the results from the first two parts, in part three of the analysis, we use simulations to investigate the estimated impact of certain changes in team heterogeneity. In general, we find that the statistically significant results are of limited economic importance in terms of differences in performances. For both the individual performance analysis as well as for the team level analysis, we find that a one standard deviation difference from the mean value, only results in performance differences well below one percent. This means that team heterogeneity is not very relevant for performance, at least within Dutch professional football.

Although the set-up of our analyses is not particularly new, our study adds to the existing literature in multiple ways. First, we use a different way of measuring performance, both at the individual level as well as at the team level. The use of objective player-match specific data is not innovative in itself within this stream of research. For example, Franck and Nüesch (2010) use this type of data to obtain a player fixed effect (talent measure) that should be constant over time. This value is then used as an input for the calculation of a coefficient of variation of team ability, which, in turn, is used in a club-match analysis of the relation between diversity and team performance. However, using a continuous Success Ratio of actions as a performance measure is new. This holds for both, the individual level as well as on an aggregated level for team performance. Besides that, our results provide new insights in the relationship between team heterogeneity and performance. It also adds to the discussion of how these variables should be measured. Furthermore, most studies only focus on a single type of heterogeneity. However, there is no *a priori* reason why different types should be studied separately. Therefore, we include four of them. Finally, our simulations reveal the economic impact of changes in heterogeneity. This is something hardly ever done in previous studies.

The rest of this paper is organized as follows. Section 5.2 provides more details on the data and the set-up of the analysis. Next, section 5.3 presents and discusses the parameter estimates. We present and discuss the results of our simulations in section 5.4. Finally, section 5.5 concludes.

5.2 Data and set-up of the analysis

We use data from the season 2014/15 of the highest tier of Dutch professional football. In this season, eighteen clubs were competing in a double round-robin competition, resulting in a total of 306 matches. At the end of the season, four clubs participated in play-off matches for qualification for UEFA Europa League football and two clubs participated in play-off matches for promotion/relegation. These matches are not included in our data. More detailed information on Dutch professional football is presented in Besters, Van Ours and Van Tuijl (2017). In total, the dataset contains 8,320 individual player-match observations for which playing time is more than zero and, accordingly, a performance measure is available. Furthermore, our dataset contains match-day related information. This includes the date and location of the match, the final outcome, the competing teams and their home provinces, whether the home team is playing on artificial grass, the attendance, general weather conditions during the match-day, fixed bookmaker odds and the pre-match rank of the competing teams. For each individual player, we have information on the date of birth, market value, height (not available for nine of the 8,320 player-match observations), nationality, the number of caps and previous playing experience in domestic leagues. Furthermore, for each player-match observation, we have playing position, playing-time, including the minute of a substitution, and whether and when someone received a first caution (i.e. a yellow card) or was expelled (i.e. a second yellow and, or red card).

From this information, we construct the set of variables as listed in panel A of Table E1 in the appendix. The variables can be subdivided in measures of *match* characteristics, *player* characteristics and *team* characteristics. Furthermore, we use Success Ratio as the measure of individual performance (dependent variable). This is the ratio of all successful actions by an individual player and all actions during a match of this individual player. The measure is appealing since it combines all sorts of actions, such as passing and duels, which makes it comparable between individuals and across playing positions. A potential downside is that no distinction is made between the importance of certain actions, i.e. all actions are equally weighted. As an alternative measure of individual performance, we use Passing%. This is the passing accuracy, measured by the number of successful passes of a certain player divided by all passes of this player during a match. In general, the remarks on the Success Ratio also apply for this performance measure. Descriptive statistics are provided in panel A of Table E2 in the appendix, while pairwise correlations between the dependent variables and all other variables are given in panel A of Table E3. Furthermore, Figure 5.1 plots two Kernel Densities. It follows that for both dependent variables, there is a high concentration of observations towards the upper bound of the scale. Although both graphs indicate a continuous pattern, the one for Passing% reveals a higher spread. Note that the statistics in the abovementioned panels as well as the graphs in Figure 5.1 are based on all player-match observations with at least 45 minutes of playing time. With this restriction, the number of observations reduces to 6,723. Since Height is missing for four observations within this subset of player-match records, the number of observations that is used in our baseline analysis is 6,719. We restrict the sample in this way, since individual player performances suffer from a lot of variation when the number of playing minutes is low. As a sensitivity analysis, we also use subsamples of players that at least played more than 0 minutes, at least played more than 15 minutes, at least played more than 30 minutes and appeared on the field as a player for 90 minutes (i.e. the whole match).

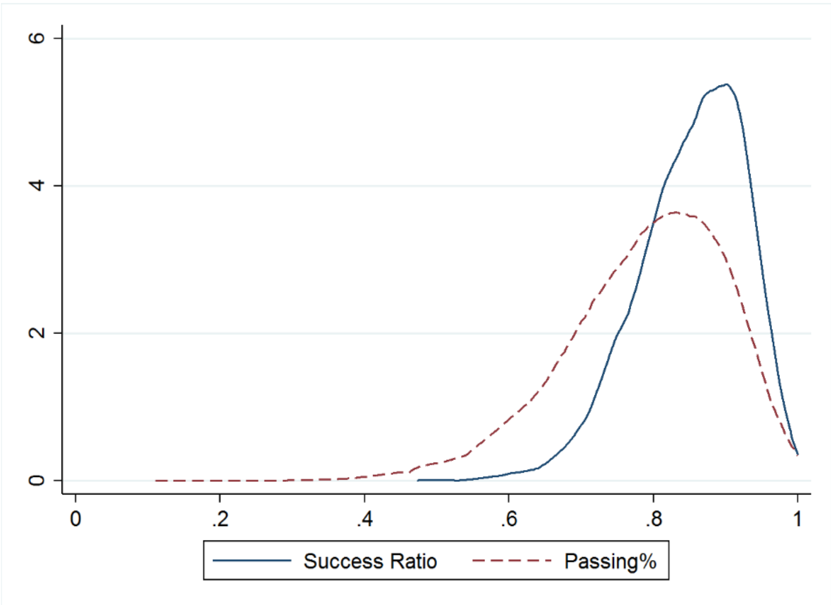


Figure 5.1: Kernel Density of individual performance measures; playing time ≥ 45 minutes

To find out the relation between individual performance and team heterogeneity, we test the following linear model with the individual player performance as the dependent variable:

$$P_{icm} = \eta_c + \lambda X_{cm} + \theta Y_{icm} + \tau Z_{icm} + \varepsilon_{icm} \qquad (1)$$

in which P_{icm} is the performance of individual i, playing for club c in match m. The vector X_{cm} represents *match* characteristics, Y_{icm} contains *player* characteristics, while Z_{icm} is the vector of *team* variables. Furthermore, λ, θ and τ are vectors of coefficients, η_c represents club fixed effects, while ε_{icm} is the error term. The vector X_{cm} contains the *match* characteristics. This includes a set of dummies for home matches, whether the match is played on artificial grass, whether the opponent is a local rival (i.e. a derby) and whether the match is played on a weekday (with Friday belonging to the weekend). Furthermore, we include as control variables the logarithm of stadium attendance and general weather conditions, measured by the temperature and by the amount of precipitation. It seems reasonable that performance also depend on the quality of the opponent. Therefore, we include a measure that captures this strength. We use relative strength measured by the expected number of points for a club (with probabilities based on bookmaker odds). A higher number of expected points suggests a weaker opponent. One could argue that this measures quality in a crude way, in particular, since it contains an average indication of the opponent. A preferred quality indicator contains an assessment of individual players that can be matched to players of the other team, for example based on playing position. Unfortunately, such information is not available. Thus, the expected number of points seems a useful approximation of the *ex-ante* relative quality of a team. Since probabilities are based on bookmaker odds, we assume that the relative strength incorporates factors such as home advantage. Furthermore, it means that quality is measured in a continuous way, thus accounting for a continuum of quality differences between clubs. As an alternative measure of relative strength, we also calculate the difference in pre-match league ranks between clubs. Finally, we use a dummy that defines whether the team is being coached by a new manager, identified as not being the one that started the season. Note that we do not include the coach, and all other (support) staff members, in our *team* characteristics and the heterogeneity measures. We suggest that on-field performance highly depends on the players on the pitch, while the influence of coaches is limited during the match. The coach can have an impact on the general team tactics, which we assume to be constant during the season and are included in the club fixed effect. The dummy for in-season coach changes captures any change in these general team-specific tactics. With respect to the differences between coaches, note that only one foreign coach has been active during the 2014/15 season.[3] Furthermore, although the experience as players can be quite different between the coaches, their experience as a club coach generally does no differ much.

---

[3] The Serb Nebojsa Gudelj, who already lived for more than fifteen years in the Netherlands, started the season as head coach of NAC Breda. He was replaced during the season by a Dutch coach.

The *player* characteristics are included in the vector Y_{icm} and relate to the individual i. We use controls for playing time, playing position dummies for goalkeeper, defender and midfielder, with attacker as the reference category, and a dummy that indicates whether someone has ever played for his national team.[4] To proxy for the quality of an individual, we use the market value of the player as given by transfermarkt.de. Furthermore, we could control for age with a continuous variable. However, we will use dummies that represent age categories for players who are aged under 21, are aged between 21 and 24, are aged between 24 and 27 and are aged over 27 (being the reference category). The cut-off ages 21, 24 and 27 are (approximately) equal to the respective 25th percentile, 50th percentile and 75th percentile of the age distribution. We also control, with a dummy, for players with a Dutch nationality and the height of the player. Finally, we include the experience of the player by the total number of matches played in domestic professional leagues. Data on matches in domestic leagues are the only reliable information on a player's entire career that we could find (this data is based on soccerway.com). Although this way of measuring match experience does not value any experience in a qualitative way, it correlates with age and seems to proxy experience quite well. In that respect, we use it as an alternative for experience measured by age.[5]

The *team* characteristics are included in the vector Z_{icm}. This vector contains a dummy that indicates whether a team member was sent off by the referee, controlling for the effect of playing with one less player.[6] Furthermore, the vector includes the weighted average values (AV) for Market Value, Height, Age and Match Experience. Weights are based on the number of playing minutes that players are together on the field. Thus, these averages (and more in general, all *team* variables) are calculated from values of team members. Note that the value for individual i is not included in the calculation of the values for these *team* characteristics. The AV variables capture any 'first moment' effect. For the effect of team heterogeneity we use, in line with previous studies, the coefficient of variation (CV). This measures diversity in Market Value, Height, Age and Match Experience. It is calculated

---

[4] Note that playing position for an individual is fixed within a match, but not between matches. Thus, for a certain match, if someone starts as a defender, he remains a defender in our analyses. However, in the next match, the same person may play as a midfielder and is counted as such for that entire match.

[5] There are multiple reasons why our measure of match experience may incorrectly measure 'true experience'. First, it does not distinguish between the level (or quality) of the league in which a match was played. Second, it does not include domestic cup and international club competition matches (e.g. Champions League). Unfortunately, reliable historical data on these types of matches is hard to find. Third, caps are not included. Fourth, international experience is valued in the same way as domestic experience. This relates to the experience of Dutch players playing domestic league matches outside of the Netherlands and similarly for other nationalities. All four aspects result in a measurement error of 'true experience'. Taking this into account, we think that, as long as we use it as an alternative for age, this should not be too much of a problem within our analyses.

[6] There is one match in which two players of the same team were sent off. However, the second player was sent off in the last minute of the match, thus it does not seem to be necessary to specify this in a different way.

by the ratio of a weighted standard deviation and a weighted mean. Weights are again based on the number of playing minutes that players are together on the field. An CV of a variable can be interpreted as the spread given its mean. In our case, a *ceteris paribus* increase in the spread means an increase in diversity. A positive coefficient means that heterogeneity is beneficial for performance, while a negative coefficient indicates that heterogeneity is bad for performance. In line with Kahane, Longley and Simmons (2013), we measure diversity in nationality with the Herfindahl-Hirschman Index (HHI). This is the sum of squared shares of nationalities, weighted by playing time that players are together on the field. If all team members of an individual i have the same nationality, the HHI is equal to one. In the other extreme that all ten team members have a different nationality, the HHI is equal to 0.1 (assuming no substitutes or sending offs, thus equal weights). An increase in diversity is reflected in a lower value of the HHI. Accordingly, a negative coefficient means that team heterogeneity is beneficial for performance. Obviously, a positive coefficient indicates the opposite.

Equation (1) also contains club fixed effects represented by η_c. These fixed effects control for unobserved heterogeneity between clubs, such as club-culture, fan base, media attention and certain club-specific aspects, such as playing tactics and transfer policies that define the composition of the squad. Thus, they control for the fact that certain teams might have younger players than other teams. The same applies to experience, market value, height and nationality. The use of these club effects means that we are explaining the variation in performance of an individual player by variation in variables that only differ between matches of the same club and a small amount of variation that results from substitutes during the match. Although one may suspect that this results in too little variation to explain differences in performance, Figures E1-E5 and Table E4-E8 in the appendix show that, within each club, there is enough variation of the relevant variables. Therefore, it seems reasonable to include the club fixed effect. Even more so, since results differ if we do not include these effects, as we shall see in the next section.

First, we discuss the set-up of the team level analyses. Since each club plays 34 matches, we are left with 612 club-match observations. The empirical approach is quite similar to the one for individuals. We now use the linear model:

$$P_{cm} = \alpha_c + \beta A_{cm} + \delta B_{cm} + \varepsilon_{cm} \qquad (2)$$

in which P_{cm} is the team performance for club c in match m. The vector A_{cm} contains *match* characteristics, while B_{cm} represents *team* characteristics. Correspondingly, β and δ are vectors of coefficients. Furthermore, α_c represent club fixed effects and ε_{cm} is the error term. Note that this model does not include any individual *player* characteristics, since we are focusing on the club level now.

For team performance, we use five different measurements. Two of them are based on the individual performance measures Success Ratio and Passing%. We calculate a weighted average, Team AV Success Ratio and Team AV Passing%, in which performances of all individual team members are weighted by playing minutes. Figure 5.2 provides Kernel Densities for both variables. The graphs show continuous distributions with most observations somewhere on the upper end of the scale. Furthermore, it reveals that there is more variability for Team AV Passing% than for Team AV Success Ratio. Panel B of Table E3 in the appendix shows that the correlation between these two variables is approximately 0.61. From that table, it also follows that these unconventional team performance measures are significantly correlated with standard measures of team performance, namely (the number of) Points, Goal Difference and Victory.[7] We discuss the differences in results between performance measures in the next section, when we present our results.

The *match* characteristics (represented by A_{cm}) are similar to those for the individual player analysis. The *team* characteristics (B_{cm}) contain the same kind of variables as for the individual players, including the team heterogeneity measures. However, these are all constructed in a slightly different way, since an average is calculated over all individuals, with weights that are based on individual playing minutes. Thus, all players of a team are used for the construction of the average. For the individual player variables we excluded the individual player i. The interpretation of the variables is unchanged. As with the individual player analysis, club fixed effects α_c control for unobserved heterogeneity between clubs. A description of the variables, as well as some descriptive statistics and the pairwise correlations are presented in the panels B of Tables E1, E2 and E3 in the appendix.

[7] Note that for the variable Points, we use two points for a win instead of three, since that makes the interpretation of our parameter estimates easier. Furthermore, Victory is a dummy with the value of one in case a team won the match and zero otherwise.

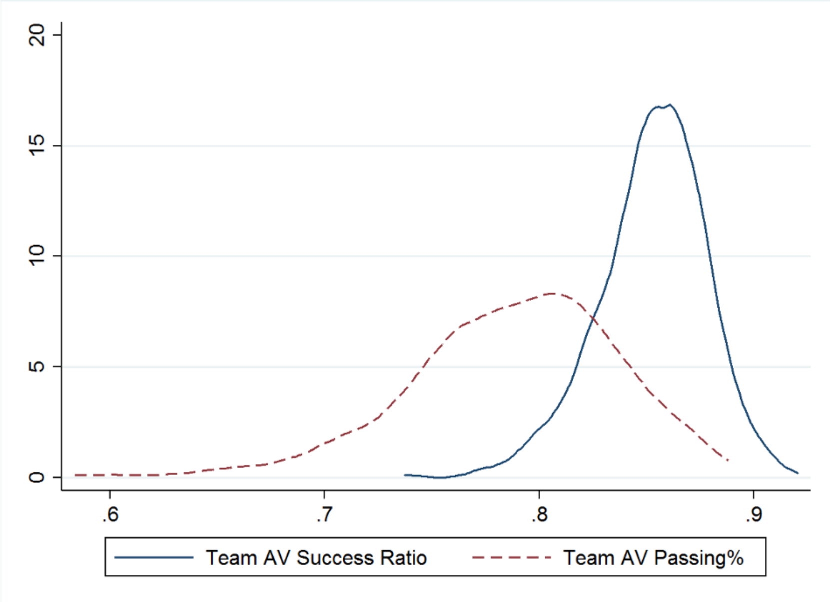


Figure 5.2: Kernel Density of team performance measured by average values of individual performances

5.3 Parameter estimates

5.3.1 Player level analysis

The baseline estimates for individual player performances are shown in Table 5.1. We distinguish between five different specifications. In all specifications, the dependent variable is the individual Success Ratio.[8] Furthermore, each specification contains 6,719 player-match observations for which playing time is at least 45 minutes. In model (1), we only include the *match* and *player* characteristics. For the *match* characteristics, it follows that, on average, individual performances are better in home matches, given the positive yet marginally significant result for the dummy. Matches that are played on artificial grass and on a weekday are insignificantly related to performance. This is also the case for derbies and the number of stadium attendance. In contrast, the general weather conditions do matter for players' successfulness of actions. Success Ratio is positively related to the temperature and negatively related to the amount of precipitation. Assuming that the higher the temperature and the lower the amount of precipitation, the better general weather conditions are. Then, our results suggest that performances are better if the weather is fine. It seems reasonable to assume that both variables capture multiple aspects that are related to these weather conditions. Such aspects include, for example, the conditions of the playing field, the mindset of the players and the physical abilities of the players. The individual performances are positively and significantly related to the expected number of points, which reflects the strength of the competing clubs. In case bookmakers expect a

[8] Table F1 in the appendix provides the results with Passing% as dependent variable. Our main results are similar.

team to do relatively well (i.e. to be stronger) compared to the opponent, players perform better. As a final *match* characteristic, we find an insignificant result for the New Coach dummy.

Table 5.1: Parameter estimates baseline results individual Success Ratio

		(1)	(2)	(3)	(4)	(5)
Match	Home	0.005*	0.005**	0.005*	0.005*	0.003*
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	Artificial Grass	0.003	0.002	0.002	0.002	0.001
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	Derby	0.004	0.004	0.004	0.004	0.002
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	Weekday	-0.004	-0.004	-0.004	-0.004	-0.003
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
	LogAttendance	-0.001	-0.002	-0.002	-0.002	0.000
		(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
	Temperature	0.006***	0.007***	0.007***	0.006***	0.006***
		(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
	Precipitation	-0.005*	-0.005*	-0.005*	-0.005*	-0.006***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	Expected Points	0.006**	0.006*	0.006*	0.006*	0.011***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	New Coach	0.003	0.004	0.004		
		(0.003)	(0.003)	(0.003)		
Player	Playing Tim/90	0.041***	0.041***	0.041***		
		(0.009)	(0.008)	(0.008)		
	Goalkeeper	0.132***	0.131***	0.132***	0.135***	0.135***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.002)
	Defender	0.094***	0.094***	0.094***	0.096***	0.096***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	Midfielder	0.048***	0.048***	0.048***	0.050***	0.050***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	Capped	0.007	0.006	0.006	0.006	0.006***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.002)
	Market Value	-0.001	-0.001	-0.001	-0.000	-0.000
		(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
	Age under 21	-0.017***	-0.016***	-0.016***	-0.016***	-0.017***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.002)
	Age between 21 and 24	-0.012***	-0.012***	-0.011***	-0.011***	-0.014***
		(0.003)	(0.003)	(0.004)	(0.004)	(0.002)
	Age between 24 and 27	-0.005	-0.004	-0.004	-0.003	-0.006***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
	Dutch	-0.005	-0.006	-0.006	-0.006	-0.004**
		(0.004)	(0.003)	(0.004)	(0.004)	(0.002)
	Height	0.063***	0.062***	0.061**	0.064**	0.050***
		(0.020)	(0.021)	(0.022)	(0.023)	(0.012)
Team	Team Member Red Card		-0.001	-0.002	-0.002	-0.003
			(0.003)	(0.003)	(0.003)	(0.003)
	AV Market Value		0.007*	0.007**	0.007**	0.005***
			(0.004)	(0.003)	(0.003)	(0.001)
	CV Market Value		-0.001			
			(0.010)			
	HHI		-0.017***	-0.015***	-0.014**	0.009***
			(0.005)	(0.005)	(0.005)	(0.003)
	AV Height		-0.025	-0.013	-0.022	-0.167***
			(0.087)	(0.083)	(0.081)	(0.050)
	CV Height		-0.343			
			(0.243)			
	AV Match Experience		-0.010	-0.010	-0.011	0.003
			(0.007)	(0.007)	(0.007)	(0.003)
	CV Match Experience		-0.017***	-0.017***	-0.017***	0.001
			(0.005)	(0.006)	(0.006)	(0.004)
	Constant	0.643***	0.732***	0.696***	0.747***	0.975***
		(0.052)	(0.184)	(0.179)	(0.178)	(0.100)
	Club Effects	Yes	Yes	Yes	Yes	No

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.* All estimates are based on 6,719 observations with playing time ≥ 45 minutes. Dependent variable is individual Success Ratio. Estimates in models (1), (2), (3) and (4) contain 18 club fixed effects with standard errors clustered by club, estimate in model (5) does not include any fixed effects.

For the *player* characteristics, we find a positive and significant result for Playing Time, even in the subgroup of player-match observations with at least 45 minutes of playing time. We suggest that there are multiple reasons for this finding. First, this might result from tactical considerations and, in particular, from the riskiness of actions that players undertake. In general, attacking (i.e. trying to score goals) is related to risky actions. Defending (i.e. preventing the other team to score), on the other hand, is less risky. Towards the end of the match, when the final outcome is quite clear for both teams, players might reduce the riskiness of actions, since the expected additional benefits are low. A second reason might be that a player's performance improves during the match, because they get used to the circumstances and the opponent. Finally, the positive relation that we find may result from poor performers that are substituted. Then, playing time is endogenous and the interpretation of the result is more difficult. The results for the position dummies are clear and highly significant. The coefficients suggest that goalkeepers perform better than defenders, who perform better that midfielders, who in turn perform better than attackers. These results are in line with the abovementioned argument of the riskiness of actions. Think, for example, of a simple pass from one central defender to the other, which is easy and without much risk. In contrast, an attacking key action in the box of the opponent is more difficult, riskier and with a higher probability of failure. The results for capped players as well as the individual market value are insignificant. Both provide some indication of the quality of the player, which does not seem to matter for individual performance. Although one might expect a positive relation between the quality of a player and his performance, we have two potential reasons for this insignificant result. First, our proxy for quality is correlated with other variables that capture part of the aspect of quality, such as position, age and experience. Second, again the riskiness of actions might play a role. If the better players 'take their responsibility', which may include undertaking certain key actions, then they will probably often fail as well. This reduces their Success Ratio. The results for the age-related dummies mean that younger players perform worse compared to older players. Furthermore, the coefficients suggest a nonlinear relation at ages below 24. The dummy for the age category between 24 and 27 is insignificant, meaning that no difference exists between this category and the reference group of players aged above 27. Although the coefficient for the dummy Dutch is negative, suggesting that Dutch players perform worse than non-Dutch players, it is insignificant. Finally, the variable Height is positive and highly significant, meaning that, on average and conditional on certain aspects, taller players perform better. This result may reflect some physical superiority in duels and, in particular, headings.

In model (2), we include the *team* characteristics. None of the results for the *match* and *player* variables is affected by this. The results for the *team* variables reveal that a red card for a team member does not affect an individual's performance, given the insignificant coefficient. Furthermore, the quality of team members, measured by their average market value, has a positive effect and is significant, tough only at the 10%-level. However, the variable CV Market Value that measures the heterogeneity in ability of teammates, is insignificant. Thus, although we find that an individual's quality does not matter for performance, the quality of team members does. Diversity in ability, however, has no effect on individual performances. For the HHI, the variable that measures the diversity in nationality of the team, we find a highly significant coefficient. The negative sign indicates that diversity is beneficial for performance. Unrelated to an individual's performance are AV Height and CV Height that measure the physical attributes of team members. Thus, we find that an individual's height matters, but the height of teammates does not. To proxy for age/experience we have age-related variables and variables that include the past playing history of players. A model specification that includes both reveals that Match Experience of teammates is preferred with Success Ratio as dependent variable, while for Passing% Age is preferred. Although the variable AV Match Experience is insignificant, the coefficient of variation of match experience is highly significant. The negative sign suggests that an increase in heterogeneity concerning the experience of team members, results in worse individual performances. This means that, related to experience, an optimum would be to play together with players that have a similar amount of experience. To sum up, while heterogeneity in nationality is good for performance, heterogeneity in experience is bad for individual performances. Furthermore, heterogeneity in ability and heterogeneity in height have no significant effect on performance.

This last result is the reason why we leave CV Market Value and CV Height out of the specification in model (3). This has no impact on any of the other results. Furthermore, all results remain the same if we exclude New Coach and Playing Time in specification (4). Both are potentially endogenous and, therefore, we test as a sensitivity analysis what happens if we leave them out. The remaining model (4) is our preferred specification and we use it for some additional analyses. First, we re-estimate the model without club fixed effects. The results are shown in model (5). Although most results for *match* and *player* characteristics are similar, the results for Capped, Age between 24 and 27 and Dutch differ and are statistically significant now. Furthermore, we find some differences for the *team* variables. First, AV Height appears to be negative and highly significant. Most interestingly, however, the sign for the HHI turns into positive, meaning that diversity in nationality is negative for performance, in contrast with the results for the within club analysis. It follows from the descriptive statistics presented in Table E5, that the HHI is quite different for the different clubs. For example, the maximum value for Willem II is the minimum value for PSV. Apparently, performance responds more to the diversity in nationality within some clubs, than in others. Given the negative coefficient as found in our baseline result in model (4), the (positive) response to diversity is larger within the clubs that generally play with a lot of different nationalities, than within clubs that only contain a small amount of diversity. In model (5) we also exploit the between-club variation in the degree of diversity. The positive sign for HHI suggests that players from clubs with little diversity in nationality perform better, on average, than players from clubs with a lot of heterogeneity in nationality. However, this result is difficult to interpret, since it might follow from unobserved club characteristics, such as playing tactics. Without club effects, we do not control for these unobserved factors. A similar reasoning applies for the result for CV Match Experience, which is insignificant in model (5), suggesting that heterogeneity in experience is unrelated to individual performances.

In general, we thus find that the use of club fixed effects matters. With these fixed effects included in the model, we analyse the variation in performance of players within a certain club. Estimation without club effects means that the differences in performances of individuals from all clubs are compared. This latter method does not take account of unobserved heterogeneity between clubs, such as the club-culture and the playing tactics. Such club-specific elements might result in club-specific playing performances. An investigation of the between-club variation might be an interesting topic for future research. Within such a design, one should carefully account for club-specific factors that may correlate with the observables.

Table 5.2: Sensitivity analysis: selection of team variables for subsets of player-match observations based on playing minutes

	(1)	(2)	(3)	(4)
Team Member Red Card	-0.007**	-0.007**	-0.002	-0.003
	(0.003)	(0.003)	(0.003)	(0.004)
AV Market Value	0.002	0.006	0.004	0.004
	(0.005)	(0.004)	(0.003)	(0.003)
HHI	-0.007	-0.005	-0.015***	-0.013**
	(0.008)	(0.006)	(0.005)	(0.005)
AV Height	0.040	-0.010	-0.008	-0.096
	(0.118)	(0.081)	(0.075)	(0.091)
AV Match Experience	-0.017**	-0.015**	-0.011*	-0.013*
	(0.007)	(0.006)	(0.006)	(0.007)
CV Match Experience	-0.030***	-0.023***	-0.021***	-0.022***
	(0.008)	(0.007)	(0.005)	(0.006)
Observations	8,311	7,690	7,038	5,095
Playing time	≥0 minutes	≥15 minutes	≥30 minutes	=90 minutes

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.* Dependent variable is individual Success Ratio. All estimates contain 18 club fixed effects and all other variables from model (4) in Table 5.1. Standard errors are clustered by club.

We proceed with the specification of model (4) in our sensitivity analysis. First we test the sensitivity of our results if we use other subsets of player-match observations. In particular, we use different cut-off values for playing time. Table 5.2 presents the results for the *team* characteristics for four different subsets. Model (1) contains all player-match observations for which playing time is more than zero. The results are very similar to those of model (2), which contains observations for playing time of at least 15 minutes. In contrast to our baseline result, we find that the effect of a red card for a team member is negative and significant in these models, while the AV Market Value and the HHI are insignificant. The result for the CV Match Experience is comparable to our baseline result, but now in combination with a significant result for the AV Match Experience. If we restrict our sample to observations for which playing time is at least 30 minutes in model (3), most of the results are similar to our baseline model (4) in Table 5.1. Furthermore, these results are also comparable to those of model (4) in Table 5.2, which contains all observations for which the player was on the field during the entire match (i.e. playing time is equal to 90 minutes). Compared to our baseline results, we find that the AV Market Value is insignificant, while the AV Match Experience is significant at the 10%-level. However, for both variables, the coefficient and the standard error do not differ much from our preferred specification. Thus, in general, we find some different results if we use different subsets. The differences occur for subsets that include player-match observations with only a small number of playing minutes. For these observations, the individual performance measure can be quite crude, since the Success Ratio is based on a rather low number of actions. This seems a good reason to exclude these observations. For subsets with the restriction that playing time has to be at least 30 minutes, we find robust results.

As a second sensitivity analysis, we test whether the results for the HHI and the CV Match Experience are different for playing position, Dutch and non-Dutch players, players aged younger than 24 and players aged older than 24 and during home matches compared to away-matches. Table 5.3 shows the results for the HHI, the CV Match Experience and some interaction terms. Interestingly, only the interaction effect NL*CV Match Experience is significantly different from zero. The positive coefficient means that Dutch players, in comparison to non-Dutch players, benefit from playing in a team with a higher diversity in experience. However, the overall effect of CV Match Experience remains negative. For playing position, age, and home matches, we do not find any significant differences. Therefore, we conclude that our baseline result is generalizable across different subgroups.

Table 5.3: Sensitivity analysis: selection of interaction effects

	(1)	(2)	(3)	(4)
HHI	-0.016*	-0.013	-0.021***	-0.020***
	(0.008)	(0.009)	(0.007)	(0.007)
CV Match Experience	-0.010	-0.041***	-0.028***	-0.015
	(0.008)	(0.011)	(0.008)	(0.009)
Goalkeeper*HHI	0.013			
	(0.016)			
Defender*HHI	-0.005			
	(0.015)			
Midfielder*HHI	0.007			
	(0.011)			
Goalkeeper*CV Match Experience	-0.021			
	(0.015)			
Defender*CV Match Experience	-0.009			
	(0.011)			
Midfielder*CV Match Experience	-0.009			
	(0.014)			
NL*HHI		-0.004		
		(0.012)		
NL*CV Match Experience		0.036**		
		(0.015)		
U24*HHI			0.011	
			(0.009)	
U24*CV Match Experience			0.021	
			(0.014)	
Home*HHI				0.014
				(0.009)
Home*CV Match Experience				-0.005
				(0.009)

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1*. All estimates are based on 6,719 observations with playing time ≥ 45 minutes. Dependent variable is individual Success Ratio. All estimates contain 18 club fixed effects and all other variables from model (4) in Table 5.1. Standard errors are clustered by club.

Finally, we have performed several additional tests for the HHI variable. The significant negative coefficient in our baseline model suggests that it is optimal to play with eleven different nationalities. This seems unlikely, because such a team composition is quite uncommon, especially given the fact that the majority of the players (within the sample) is Dutch. Therefore, we look for non-linearity with dummy variables that each represent about 10 percent of the observations. The results are presented in the first column of Table F2 in the appendix. For the individual player level, we find that all coefficients are negatively signed. However, only few are significant, without a clear pattern (of non-linearity). We suggest that the within-club design has an influence on these results. Given the differences between the clubs, a dummy may only contain observations of certain specific clubs. This also follows from the spread as shown in Figure E2. Furthermore, models (3) and (5) contain alternative specifications of the HHI variable. In particular, the variable is calculated for a specific group of players that typically interact with each other during the match. Model (3) contains the separate variables for defenders, including goalkeepers, midfielders and attackers. Model (5) contains separate HHI variables for players with a position on the left-wing, players within the central axis of the team and players with a position on the right-wing. Only for the players on the left-wing, we find a significant result (at a 10 percent level). In general, we conclude that team diversity with respect to nationality does not differ between separate parts of the team. Although players within these separate parts will probably communicate and interact with each other (i.e. players communicate most of all with the players quite close to them). However, within our sample of 523 players, 434 speak Dutch, while 68 speak English. The remaining 21 players speak other languages.[9] Given that the vast majority speaks Dutch and/or English, we suggest that language cannot be considered as an important element, in particular because Dutch people, in general, speak English quite well. Instead, we think that differences in culture, background, and education in football at young ages (i.e. ‘the Dutch system’ vs. other playing styles) might be important elements that are related to the nationality of players. Furthermore, we suggest three alternative reasons why we think that language is not an issue. First, international players who move to the Netherlands are (probably) ambitious and want to move to other leagues after several seasons. Therefore, it is reasonable to assume that they want to learn English, if they not already speak this. Second, clubs will want to invest in their assets, i.e. their players. Nowadays, the clubs guide and support the foreign players and provide language courses so that communication becomes easier. Finally, we suggest that the difficulty and intensity of the language that is used during a match is not very high. Players will easily learn the basics necessary to communicate on the pitch, whatever mother language.

5.3.2 Team level analysis

The setup of our team level analyses is similar to the analyses for individual player performances. The results are shown in Table 5.4, where we distinguish between five different dependent variables. In the models (1) and (2) we use a team average of Success Ratio and a team average of Passing% as team performance indicators, respectively. These are calculated as a weighted average of individual performances with weights that are based on playing time. In the models (3), (4) and (5) we use common match-outcomes as dependent variables. Points is measured as zero for a loss, one for a draw and two for a win to make interpretation of the results easier. Furthermore, Goal Difference is the number of goals scored minus the number of goals conceded, while Victory is a dummy with the value of one in case a team has won the match. All estimates are based on 612 club-match observations and include club fixed effects. The *match* characteristics are the same as for the individual player analyses, except the measure for the relative strength of opponents. Instead of the expected number of points, we use the difference in pre-match rank. The relation between the expected number of points and actual points might reveal more about the efficiency of bookmakers

[9] This data is based on nationality, place of birth, playing history and YouTube video-interviews.

than that it measures the quality of teams. We find that team performance is better for home playing clubs in all models. Furthermore, in the models (1) and (2) we find a significant and positive result for Temperature, which is in line with our results for individual player performances. However, Temperature is insignificant in models (3), (4) and (5). In these three models, the difference in rank is negative and highly significant. Note that the league leader has the rank one and the team at the bottom of the league table has the rank eighteen. Thus, a negative value for the difference in pre-match rank means that a team is ranked higher than the opponent. The negative coefficient means that higher ranked teams perform better. The difference in rank is only marginally significant (at a 10 percent level) in model (1). Furthermore, the *match* characteristics are insignificant.

Table 5.4: Team level analysis

		(1)	(2)	(3)	(4)	(5)
	Dependent variable	Team AV Success Ratio	Team AV Passing%	Points	Goal Difference	Victory
Match	Home	0.007***	0.009**	0.274***	0.592***	0.137***
		(0.002)	(0.003)	(0.050)	(0.124)	(0.028)
	Artificial Grass	0.004	0.002	0.033	-0.059	0.065
		(0.003)	(0.007)	(0.106)	(0.277)	(0.052)
	Derby	0.003	0.007	0.031	0.126	0.059
		(0.003)	(0.007)	(0.146)	(0.255)	(0.083)
	Weekday	-0.003	0.003	0.001	-0.004	0.011
		(0.004)	(0.009)	(0.218)	(0.446)	(0.114)
	LogAttendance	-0.001	-0.003	0.026	-0.087	0.043
		(0.002)	(0.004)	(0.067)	(0.187)	(0.027)
	Temperature	0.006***	0.015***	0.033	0.082	0.010
		(0.001)	(0.003)	(0.060)	(0.141)	(0.027)
	Precipitation	-0.004	-0.011	-0.005	0.011	-0.019
		(0.002)	(0.007)	(0.077)	(0.165)	(0.047)
	Difference in Rank	-0.000*	0.000	-0.066***	-0.130***	-0.032***
		(0.000)	(0.000)	(0.005)	(0.011)	(0.003)
Team	Red Card in Team	-0.002	-0.013**	-0.218***	-0.729***	-0.086
		(0.002)	(0.005)	(0.074)	(0.170)	(0.050)
	Team AV Market Value	0.006*	0.014*	-0.079	-0.192	-0.058
		(0.003)	(0.007)	(0.113)	(0.435)	(0.080)
	Team HHI	-0.017**	-0.037**	0.388	0.731	0.117
		(0.006)	(0.016)	(0.252)	(0.691)	(0.162)
	Team AV Height	0.032	0.041	-9.295***	-24.893***	-3.936**
		(0.085)	(0.235)	(2.546)	(6.696)	(1.769)
	Team AV Match Experience	-0.004	-0.015	-0.254	-0.871**	-0.075
		(0.006)	(0.011)	(0.198)	(0.408)	(0.125)
	Team CV Match Experience	-0.018**	-0.033	-0.861***	-1.827***	-0.354**
		(0.007)	(0.022)	(0.262)	(0.612)	(0.132)
	Constant	0.811***	0.771*	18.319***	47.989***	7.389**
		(0.162)	(0.428)	(4.721)	(12.825)	(3.382)

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.* All estimates are based on 612 observations and contain 18 club fixed effects. Standard errors are clustered by club.

For the *team* characteristics, we find some interesting differences between the results of models (1) and (2) and the other three specifications. Although the difference for Red Card in Team is not so clear, Team AV Market Value is positive and significant (at the 10%-level) in the first two models, while negative and insignificant in the models (3)-(5). Thus, the average quality of a team, measured by market values, has no direct effect on match-outcomes, but is positively related to average team performances. A similar conclusion follows from the results for the HHI. An increase in diversity is positive for average team performance, while there is no direct effect on match-outcomes.[10] In contrast, Team AV Height has a direct and negative effect on match-outcomes, while it does not have any effect on team average performances. The insignificant result in the models (1) and (2) is not surprising and in line with our results for individual player performance. However, the direct and highly significant results in models (3)-(5) are more puzzling. This might have to do with match-specific playing tactics and the use of taller players in *ex ante* difficult matches. An alternative explanation might be that the better clubs have, on average, smaller players, with more technical skills, which is beneficial when possessing the ball. We focus on team heterogeneity and find a rather robust result for CV Match Experience. Although the coefficient is insignificant in model (2), it is negative and significant in the other specifications. This suggests a direct effect of team heterogeneity in experience on match-outcomes as well as on average team performance. In line with our result for individual performance, this means that it is better to compose a team of players with an equal amount of experience.

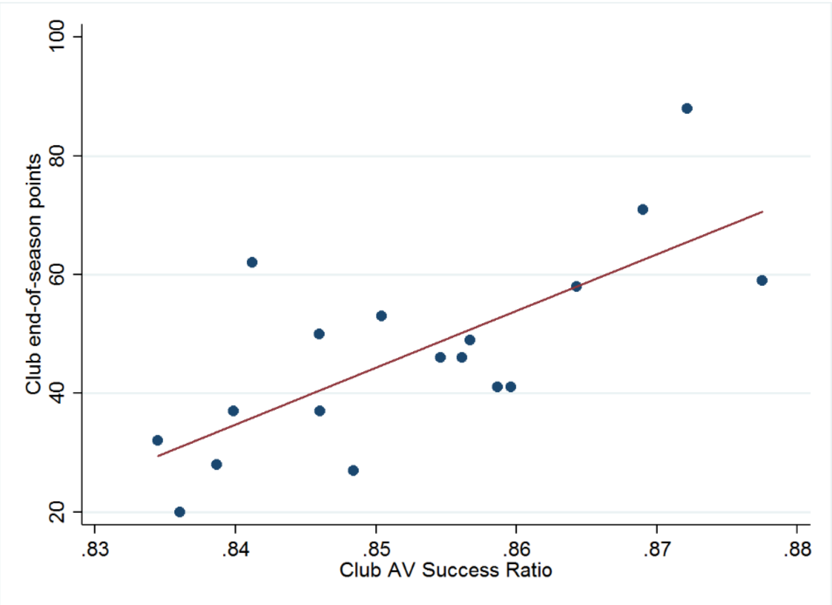


Figure 5.3: Scatterplot of club average Success Ratio and club end-of-season number of points (win 3 points, draw 1 point)

[10] The models (2), (4) and (6) in Table F2 provide the results for some variants of the HHI variable. Comparable to the results for the individual player analysis, we do not find evidence of a non-linear relationship. Furthermore, with two exceptions, we find insignificant results for the separate groups of players based on playing position.

In the above, we referred to direct effects on match-outcomes, which suggests that there might also be indirect effects. For example, something might have an effect on team average performance, measured by an average Success Ratio, which in turn is related to match-outcomes. In particular, this might be the case for the AV Market Value and the HHI. We, therefore, examine the relation between average team performances and match-outcomes in more detail. First, Figure 5.3 plots, for all eighteen clubs, the average seasonal Success Ratio and the end-of-season number of points.[11] This shows a positive relation. Furthermore, Table 5.5 shows the results for the *team* characteristics if Team AV Success Ratio is included as an independent variable and match-outcomes are the dependent variables. In all three models, the average success ratio is positive and highly significant. This indeed suggests that there are indirect effects at work. Furthermore, it also reveals the importance of the way in which team performance is measured. A continuous variable might be more accurate and able to reveal certain relations than discrete and crude match-outcomes.

Table 5.5: Team level analysis with Team AV Success Ratio as independent variable

		(1)	(2)	(3)
	Dependent variable	Points	Goal Difference	Victory
Team	Red Card in Team	-0.208**	-0.699***	-0.081
		(0.074)	(0.166)	(0.050)
	Team AV Market Value	-0.107	-0.273	-0.073
		(0.112)	(0.441)	(0.081)
	Team HHI	0.462*	0.947	0.157
		(0.238)	(0.671)	(0.158)
	Team AV Height	-9.437***	-25.310***	-4.012**
		(2.527)	(6.371)	(1.760)
	Team AV Match Experience	-0.237	-0.821**	-0.066
		(0.190)	(0.364)	(0.121)
	Team CV Match Experience	-0.781***	-1.590**	-0.311**
		(0.251)	(0.552)	(0.127)
	Team AV Success Ratio	4.372***	12.876***	2.364**
		(1.461)	(3.143)	(0.912)

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.* All estimates are based on 612 observations and contain 18 club fixed effects and all *match* characteristics included in Table 5.4. Standard errors are clustered by club.

5.4 Simulations

In the previous sections, we discussed a preferred model specification and the statistical significance of variables. In this section, we investigate the economic impact and relevance of our main results. We use model simulations to see how performance changes for different values of the AV Market Value, the HHI, the AV Height and the CV Match Experience. First, we look at the impact for individual players and use the results of model (4) in Table 5.1. In all these simulations we set the Home dummy at value one, while Artificial Grass, Derby and Weekday are set to zero. Furthermore, we assume the player to be a Dutch midfielder aged between 24 and 27. Thus, the dummies

[11] Figure G1 in the appendix shows the same relation in case we use two points for a win instead of three.

Midfielder, Dutch and Age between 24 and 27 have a value of one, while the other position and age-related dummies have a value of zero. We also assume that the player has never played for his national team, i.e. Capped has a value of zero. Team Member Red Card is set to zero as well. All other variables are set at sample mean. First, for the four variables of interest, we add (subtract) one standard deviation to (from) its sample mean. Next, we use the minimum and maximum values in our simulations. Table 5.6 shows the results. The last column reveals that none of the simulated values results in a difference that is (even) close to one percent.

Table 5.6: Simulation of individual player performance

		Value	Success Ratio	Absolute difference compared to mean	% difference compared to mean
AV Market Value	Min	0.14	0.8366	-0.0075	-0.0089
	Mean-SD	0.15	0.8367	-0.0074	-0.0088
	Mean	1.25	0.8441	0	0
	Mean+SD	2.35	0.8515	0.0074	0.0088
	Max	5.18	0.8707	0.0266	0.0315
HHI	Min	0.13	0.8494	0.0053	0.0063
	Mean-SD	0.29	0.8471	0.0030	0.0036
	Mean	0.50	0.8441	0	0
	Mean+SD	0.71	0.8411	-0.0030	-0.0036
	Max	1.00	0.8370	-0.0071	-0.0085
AV Height	Min	1.77	0.8452	0.0011	0.0013
	Mean-SD	1.80	0.8445	0.0004	0.0005
	Mean	1.82	0.8441	0	0
	Mean+SD	1.84	0.8437	-0.0004	-0.0005
	Max	1.86	0.8432	-0.0009	-0.0010
CV Match Experience	Min	0.36	0.8508	0.0067	0.0080
	Mean-SD	0.57	0.8472	0.0031	0.0037
	Mean	0.75	0.8441	0	0
	Mean+SD	0.93	0.8410	-0.0031	-0.0037
	Max	1.48	0.8315	-0.0126	-0.0149

Note: Simulation based on results of model (4) in Table 5.1. In all these simulations we set the Home dummy at value one, while Artificial Grass, Derby and Weekday are set to zero. Furthermore, the dummies Midfielder, Dutch and Age between 24 and 27 were set at value one, while the other position and age-related dummies were set at value zero. Also, Capped is set at value zero as well as Team Member Red Card. All other variables are set at sample mean.

A similar result appears for simulations for our team level analyses. For the *match* characteristics, the same assumptions are made as described above, while Difference in Rank is set to zero. Table 5.7 shows the simulated results for Team AV Success Ratio. Table H1 in the appendix provides the results for all other dependent variables. Note that we have no simulated value for Mean-SD for AV Market Value, since that would be less than the minimum value. Again, it follows that the differences are very small. Thus, in general, we conclude that our statistical significant results have hardly any economic relevance, at least in the setting of Dutch professional football.

Finally, related to the experience of a team, we discuss the optimal level of experience that follows from our results. Since the coefficients for the AV Match Experience variable are negative, the addition of experience to the team reduces the performance. For small changes in match experience, this effect is stronger than the effect that results from the value for the CV Match Experience, which works in the opposite direction. However, we cannot easily identify an optimal level of experience. In particular, because (most of) the coefficients for the AV Match Experience are insignificant. Based on the significant negative value for CV Match Experience, we conclude that it is better to have a team composed of players with an equal experience than with players who differ in experience. The result suggests that this can be of any level of experience.

Table 5.7: Simulation of team performance measured by Team AV Success Ratio

		Value	Team AV Success Ratio	Absolute difference compared to mean	% difference compared to mean
AV Market Value	Min	0.17	0.8482	-0.0068	-0.0080
	Mean-SD				
	Mean	1.25	0.8551	0	0
	Mean+SD	2.35	0.8620	0.0070	0.0081
	Max	4.85	0.8779	0.0228	0.0267
HHI	Min	0.14	0.8609	0.0059	0.0069
	Mean-SD	0.28	0.8586	0.0035	0.0041
	Mean	0.49	0.8551	0	0
	Mean+SD	0.70	0.8515	-0.0035	-0.0041
	Max	1.00	0.8465	-0.0085	-0.0100
AV Height	Min	1.79	0.8541	-0.0010	-0.0011
	Mean-SD	1.81	0.8547	-0.0003	-0.0004
	Mean	1.82	0.8551	0	0
	Mean+SD	1.83	0.8554	0.0003	0.0004
	Max	1.85	0.8560	0.0010	0.0011
CV Match Experience	Min	0.42	0.8613	0.0063	0.0073
	Mean-SD	0.59	0.8582	0.0031	0.0037
	Mean	0.76	0.8551	0	0
	Mean+SD	0.93	0.8519	-0.0031	-0.0037
	Max	1.34	0.8444	-0.0107	-0.0125

Note: Simulation based on results of model (1) in Table 5.4. In all these simulations we set the Home dummy at value one, while Artificial Grass, Derby and Weekday are set to zero. Also, Team Member Red Card is set at zero. All other variables are set at sample mean.

Table 5.8: Variation in experience

Club	Start Match	End Match	Difference	SD Difference	SD Difference between Matches
ADO Den Haag	138.6	129.1	9.5	14.2	11.7
AZ	88.3	92.9	-4.7	15.7	12.3
Ajax	83.6	84.7	-1.1	16.6	18.4
Excelsior	104.0	93.5	10.5	8.7	12.2
FC Dordrecht	54.0	59.9	-5.9	13.6	14.8
FC Groningen	125.3	118.4	6.8	14.4	14.3
FC Twente	93.4	94.4	-1.0	12.7	13.4
FC Utrecht	102.1	94.8	7.3	19.5	17.0
Feyenoord	134.1	116.6	17.5	26.1	23.0
Go Ahead Eagles	108.1	101.6	6.6	24.0	17.8
Heracles Almelo	113.4	113.1	0.2	20.1	21.0
NAC Breda	121.0	116.6	4.4	15.0	16.7
PEC Zwolle	115.8	117.6	-1.8	20.3	17.0
PSV	104.7	97.1	7.7	11.5	10.6
SC Cambuur	120.9	119.8	1.1	12.9	10.7
SC Heerenveen	75.1	71.3	3.9	8.2	7.8
Vitesse	118.4	118.6	-0.2	11.0	15.6
Willem II	124.8	122.1	2.7	19.5	15.0
Total	107.0	103.5	3.5	15.8	15.0

In any case, the optimal level seems to be club specific, since mean values of match experience differ between clubs. The first column of Table 5.8 provides the mean values of experience for the starting line-up. Furthermore, the second column provides mean values for the players who end the match, while the third column gives the mean difference. It follows that for some clubs, the in-match substitutions result in a reduction of experience, while for other clubs, the experience increases. On average, the value is positive, suggesting that the starting eleven are more experienced than the players who end the match. The fourth column gives the standard deviation of the differences. This can be compared to the differences that result from changes between matches. The last column shows the standard deviation of the mean values of experience of the starting line-ups for all the 34 matches. In general, the values in columns four and five are quite comparable. Thus, suggesting that coaches substitute in a similar way within matches as between matches, at least with respect to experience. Note that the differences in column three and the standard deviations in column four are not weighted by playing minutes. Thus, it simply reveals what kind of changes are made during the match, without showing the eventual impact on the values to be used in our estimations. Since about 95 percent of the substitutions take place in the second half of the match, this impact is at least half of the values provided.

5.5 Discussion and conclusion

The primary focus of this paper is to investigate the relationship between team heterogeneity and performance. We use data from the highest tier of Dutch professional football in the 2014/15 season and construct four types of heterogeneity measures. The first one is based on market values that proxy for skill and ability. Second, we construct a measure for diversity in nationality. Next, height is used as a measure of heterogeneity of physical ability, while age and past playing histories are used to proxy for diversity in experience. First, we test whether team heterogeneity has an effect on individual player performances. A linear model with an individual's Success Ratio as dependent variable is used for this. After controlling for numerous *match* and *player* characteristics, we find that heterogeneity in ability and diversity in height are unrelated to performance. However, team heterogeneity in nationality has a positive effect on performance, while team heterogeneity in experience has a negative effect. These findings are robust according to numerous sensitivity analyses and depend on the use of club fixed effects.

The result for diversity in nationality suggests that benefits overcome certain cultural barriers, i.e. integration costs. Adding non-Dutch players to a squad increases diversity, in particular if those players come from different countries. Our results do not reveal what specific elements drive the benefits. For example, these elements might be related to different educational systems and talent development programs in youth football. However, we do know that the ones that often play in the Dutch league have a positive impact on individual's performance. Assuming that this is important, one should not be pessimistic about the inflow of foreign players, at least up to the level that was reached in the 2014/15 season.

The results for heterogeneity in experience reveal that diversity is bad for performance. Players do better in case their teammates have a comparable amount of experience. We suggest two potential reasons for this. First, communication (in a broad sense) might be more difficult between players with different levels of experience. The players might suggest different solutions to certain situations and do not understand why the other players thinks differently. Second, in line with previous studies, this results from the complementarity of tasks within football. Mutual learning is of minor importance on the field. To perform well is what matters and this is achieved by combining complements in a team.

In general, our main findings are the same for a team level analyses, when a team average performance is used as dependent variable. Results are somewhat different if we use match-outcomes instead. These match-outcomes, however, are related to team average performances. Thus, we suggest that, there are direct effects as well as indirect effects of team heterogeneity on match-outcomes. These indirect effects work through team average performance, in particular for heterogeneity concerning nationality. It seems that discrete match-outcomes are not able to capture such influences. Continuous measures, such as a (team average) Success Ratio and a (team average) Passing% might overcome this problem. Although these measures have drawbacks as well, in particular in assigning similar weights to all actions, they seem useful in this type of analysis. We suggest that an improvement of the understanding of the relationship between team heterogeneity and performance can be made if more detailed (performance) data becomes available. However, it is questionable whether this will make a difference in our conclusions on the economic impact. The results of our simulations all suggests very little relevance in terms of additional performance. Of course, this result is obtained from Dutch professional football and, therefore, not necessarily the same in other situations. Future research might extend the analyses to multiple seasons and leagues from professional football. It is also interesting to look outside the domain of sports, if clear measures of individual performances are available.

Appendix E: Details on the data

The sample contains 306 regular league matches from the highest tier of Dutch professional football in the season 2014/15. In every match, the line-up of a team consists of eleven players. Furthermore, a maximum of 7 players is allowed to take place on the bench, while each team is allowed to make a maximum of three substitutes per match. For each team, there is one goalkeeper. The number of defenders, midfielders and forwards depends on the team tactics and can change during a match as a result of substitutions. A match is split in two halves, both with an official playing time of 45 minutes. This can be extended with some stoppage time, but since we do not have precise data on this, we assume playing time to be 90 minutes at maximum. This means that players who entered the match in the stoppage time, are assumed to have a playing time equal to zero. In case no individual player performance data is available, playing time is also assumed to be zero. In total, we have 8,320 player-match observations for which playing time is more than zero and, accordingly, an individual performance measure is available.

Data on line-ups (including coaches), substitutes, playing time, cards, playing position, nationalities and market values were collected from transfermarkt.de. Furthermore, we also used this source to collect data on match-day information, such as the date of the match, the pre-match rank of teams and attendance figures. Information on weather conditions comes from the KNMI (Royal Dutch Meteorological Institute) which is located in De Bilt (rather central within the Netherlands). Fixed bookmaker odds come from the betting agency William Hill and are obtained from football-data.co.uk. Birth dates of players as well as the number of caps are collected from vi.nl. Height is mainly collected from soccerway.com, but in case of missing data, other sources were consulted. We still lack information about height for nine of the 8,320 player-match observations. In case height is missing, a player is not considered in the construction of the team variables for height. Playing history, in particular in domestic league matches, is collected from soccerway.com. The player-specific match data, which includes the individual performance measures, is provided by ORTEC Sports.

From this dataset, we construct the variables as listed in Table E1. Table E2 provides some descriptive statistics. Table E3 reports pairwise correlations. In both tables, the values for player level data are based on the subsample of players with a playing time of at least 45 minutes. Furthermore, Figures E1-E5 provide densities by club for some key variables, while Tables E4-E8 show some descriptive statistics of these variables by club.

Table E1: Description of variables

	A. Individual Players	
DV	**Success Ratio**	Ratio of the number of successful action and the total number of action
	Passing%	Percentage of successful passes
Match	**Home**	Dummy with value one if club played at home
	Artificial Grass	Dummy with value one if match is played on artificial grass
	Weekday	Dummy with value one if match was played on a Monday, Tuesday, Wednesday or Thursday
	Derby	Dummy with value one if clubs are from the same province
	LogAttendance	Natural logarithm of stadium attendance
	Temperature	Daily mean temperature measured in 0.1 degrees Celsius in De Bilt on match-day / 10
	Precipitation	Daily precipitation (precipitation was measured as being -1 for values <0.05, but was set to zero in our sample) amount in 0.1 mm in De Bilt on match-day / 10
	Expected Points	Expected number of points for club based on bookmaker odds calculated as win-probability * 3 + draw-probability * 1 + lose-probability * 0
	Rank	Pre-match rank of club
	Rank Opponent	Pre-match rank of the opponent
	Difference in Rank	Pre-match rank of club minus pre-match rank of the opponent
Player	**Playing Time**	Playing time during a match in minutes / 90 (maximum is 90 minutes)
	Goalkeeper	Dummy with value one if playing position is goalkeeper
	Defender	Dummy with value one if playing positon is defender
	Midfielder	Dummy with value one if playing positon is midfielder
	Attacker	Dummy with value one if playing positon is attacker
	Capped	Dummy with value one if player at least played once for the national team
	Market value	Market value (measured the first time a player is in the team during the season; in €1,000,000)
	Age	Age at the start of the season (8th of August 2014; in days divided by 365.25)
	Age under 21	Dummy with value one if the age at start of the season <21
	Age between 21 and 24	Dummy with value one if the age at start of the season ≥21 and <24
	Age between 24 and 27	Dummy with value one if the age at start of the season ≥24 and <27
	Age above 27	Dummy with value one if the age at start of the season ≥27
	Dutch	Dummy with value one if player has the Dutch nationality (only first nationality taken into account)
	Height	Height of player in cm / 100
	Match Experience	Number of matches played in professional football in domestic leagues prior to the start of the 2014/15 season on 8 August 2014 (divided by 100)
Team	**Team Member Red Card**	Dummy with value one if at least one team member was sent off by the referee
	AV Market Value	Weighted average market value of team members. Weights are based on playing time that players were together on the field
	SD Market value	Weighted standard deviation of market value of team members. Weights are based on playing time that players were together on the field
	CV Market value	Coefficient of variation of market value calculated as SD Market value / AV Market value
	HHI	Weighted Herfindahl-Hirschman Index of player nationalities of team members. Weights are based on playing time that players were together on the field
	AV Height	Weighted average height of team members. Weights are based on playing time that players were together on the field
	SD Height	Weighted standard deviation of height of team members. Weights are based on playing time that players were together on the field
	CV Height	Coefficient of variation of height calculated as SD Height / AV Height
	AV Age	Weighted average age of team members. Weights are based on playing time that players were together on the field
	SD Age	Weighted standard deviation of age of team members. Weights are based on playing time that players were together on the field
	CV Age	Coefficient of variation of match experience calculated as SD Age / AV Age
	AV Match Experience	Weighted average match experience of team members. Weights are based on playing time that players were together on the field
	SD Match Experience	Weighted standard deviation of match experience of team members. Weights are based on playing time that players were together on the field
	CV Match Experience	Coefficient of variation of age calculated as SD Match Experience / AV Match Experience

	B. Team Level	
DV	**Team AV Success Ratio**	Weighted average Success Ratio. Weights are based on playing time
	Team AV Passing Accuracy	Weighted average Passing Accuracy. Weights are based on playing time
	Points	Number of points obtained in the match
	Goal difference	Goals scored minus goals conceded
	Victory	Dummy with value one if match is won
Team	**Team AV Market Value**	Weighted average Market Value of team. Weights are based on playing time
	Team SD Market Value	Weighted standard deviation of Market Value of team. Weights are based on playing time
	Team CV Market Value	Coefficient of variation of Market Value of team calculated as Team SD Market Value / Team AV Market Value
	Team HHI	Weighted Herfindahl-Hirschman Index of player nationalities of team. Weights are based on playing time
	Team AV Height	Weighted average height of team. Weights are based on playing time
	Team SD Height	Weighted standard deviation of height of team. Weights are based on playing time
	Team CV Height	Coefficient of variation of height of team calculated as Team SD Height / Team AV Height
	Team AV Age	Weighted average age of team. Weights are based on playing time
	Team SD Age	Weighted standard deviation of age of team. Weights are based on playing time
	Team CV Age	Coefficient of variation of age of team calculated as Team SD Age / Team AV Age
	Team AV Match Experience	Weighted average Match Experience of team. Weights are based on playing time
	Team SD Match Experience	Weighted standard deviation of Match Experience of team. Weights are based on playing time
	Team CV Match Experience	Coefficient of variation of Match Experience of team calculated as Team SD Match Experience / Team AV Match Experience

Note: DV is dependent variable

Table E2: Descriptive statistics

	A. Player Level	N	Mean	SD	Median	Min	Max
DV	Success Ratio	6,723	0.85	0.07	0.86	0.47	1.00
	Passing%	6,723	0.79	0.11	0.80	0.11	1.00
Match	Home	6,723	0.50	0.50	0	0	1
	Artificial Grass	6,723	0.33	0.47	0	0	1
	Derby	6,723	0.11	0.31	0	0	1
	Weekday	6,723	0.04	0.19	0	0	1
	LogAttendance	6,723	9.60	0.71	9.63	8.05	10.87
	Temperature	6,723	0.91	0.54	0.90	-0.28	1.91
	Precipitation	6,723	0.17	0.31	0.00	0	2.00
	Expected Points	6,723	1.38	0.55	1.35	0.19	2.73
	Difference in Rank	6,723	0	7.80	1	-17	17
Player	Playing Time/90	6,723	0.95	0.11	1.00	0.50	1.00
	Goalkeeper	6,723	0.09	0.29	0	0	1
	Defender	6,723	0.36	0.48	0	0	1
	Midfielder	6,723	0.27	0.45	0	0	1
	Attacker	6,723	0.27	0.44	0	0	1
	Capped	6,723	0.25	0.43	0	0	1
	Market Value	6,723	1.26	1.52	0.60	0.00	8.00
	Age	6,723	24.08	3.42	23.74	16.52	36.67
	Age under 21	6,723	0.19	0.39	0	0	1
	Age between 21 and 24	6,723	0.35	0.48	0	0	1
	Age between 24 and 27	6,723	0.26	0.44	0	0	1
	Age above 27	6,723	0.20	0.40	0	0	1
	Dutch	6,723	0.64	0.48	1	0	1
	Height	6,719	1.82	0.06	1.82	1.63	1.99
	Match Experience	6,723	1.07	0.83	0.90	0	5.26
Team	Team Member Red Card	6,723	0.08	0.27	0.00	0	1
	AV Market Value	6,723	1.25	1.10	0.80	0.14	5.18
	CV Market Value	6,723	0.65	0.18	0.62	0.17	1.47
	HHI	6,723	0.50	0.21	0.50	0.13	1.00
	AV Height	6,723	1.82	0.02	1.82	1.77	1.86
	CV Height	6,723	0.03	0.01	0.03	0.01	0.05
	AV Age	6,723	24.07	1.26	24.15	20.34	27.83
	CV Age	6,723	0.13	0.03	0.13	0.05	0.24
	AV Match Experience	6,723	1.06	0.26	1.08	0.25	1.83
	CV Match Experience	6,723	0.75	0.18	0.74	0.36	1.48
	B. Team Level	N	Mean	SD	Median	Min	Max
DV	Team AV Success Ratio	612	0.85	0.02	0.86	0.74	0.92
	Team AV Passing%	612	0.79	0.05	0.79	0.58	0.89
	Points	612	1	0.87	1	0	2
	Goal Difference	612	0	1.94	0	-5	5
	Victory	612	0.38	0.49	0	0	1
Match	Home	612	0.50	0.50	0.50	0	1
	Artificial Grass	612	0.33	0.47	0	0	1
	Derby	612	0.11	0.31	0	0	1
	Weekday	612	0.04	0.19	0	0	1
	LogAttendance	612	9.60	0.71	9.64	8.05	10.87
	Temperature	612	0.91	0.54	0.90	-0.28	1.91
	Precipitation	612	0.17	0.31	0	0	2
	Expected Points	612	1.38	0.55	1.35	0.19	2.73
	Difference in Rank	612	0	7.80	0	-17	17
Team	Red Card in Team	612	0.09	0.29	0	0	1
	Team AV Market Value	612	1.25	1.10	0.81	0.17	4.85
	Team CV Market Value	612	0.65	0.17	0.62	0.22	1.24
	Team HHI	612	0.49	0.21	0.49	0.14	1
	Team AV Height	612	1.82	0.01	1.82	1.79	1.85
	Team CV Height	612	0.03	0.01	0.04	0.02	0.05
	Team AV Age	612	24.07	1.22	24.14	20.79	27.15
	Team CV Age	612	0.13	0.03	0.13	0.07	0.22
	Team AV Match Experience	612	1.06	0.25	1.08	0.32	1.68
	Team CV Match Experience	612	0.76	0.17	0.75	0.42	1.34

Note: For Player Level, selection of observations with playing time of at least 45 minutes. Four observations are missing for Height because of missing data. DV is dependent variable.

Table E3: Pairwise correlations

	A. Player Level	Success Ratio	Passing%
DV	Success Ratio	1	
	Passing%	0.5775*	1
Match	Home	0.0565*	0.0499*
	Artificial Grass	-0.0127	-0.0311*
	Derby	0.0045	-0.0031
	Weekday	-0.0098	0.0005
	LogAttendance	0.0324*	0.0550*
	Temperature	0.0411*	0.0555*
	Precipitation	-0.0248*	-0.0428*
	Expected Points	0.1357*	0.1473*
	Difference in Rank	-0.1075*	-0.0976*
Player	Playing Time/90	0.2236*	0.1037*
	Goalkeeper	0.3410*	0.0258*
	Defender	0.3681*	0.2555*
	Midfielder	-0.1050*	0.0457*
	Attacker	-0.5138*	-0.3391*
	Capped	0.0757*	0.1074*
	Market Value	0.0695*	0.1052*
	Age	0.1692*	0.0789*
	Age under 21	-0.1141*	-0.0417*
	Age between 21 and 24	-0.0753*	-0.0704*
	Age between 24 and 27	0.0715*	0.0537*
	Age above 27	0.1221*	0.0652*
	Dutch	0.0482*	0.014
	Height	0.2373*	0.1286*
	Match Experience	0.1219*	0.0866*
Team	Team Member Red Card	-0.018	-0.0377*
	AV Market Value	0.1269*	0.1610*
	CV Market Value	0.0179	0.0203
	HHI	-0.0096	-0.015
	AV Height	-0.1643*	-0.1460*
	CV Height	-0.0171	0.0041
	AV Age	-0.0620*	-0.0577*
	CV Age	0.005	0.0071
	AV Match Experience	0.0139	0.0203
	CV Match Experience	-0.0358*	-0.0377*

	B. Team Level	Team AV Success Ratio	Team AV Passing%	Points	Goal Difference	Victory
DV	Team AV Success Ratio	1				
	Team AV Passing%	0.6121*	1			
	Points	0.2946*	0.1118*	1		
	Goal Difference	0.3376*	0.1420*	0.8786*	1	
	Victory	0.2799*	0.1307*	0.8985*	0.7895*	1
Match	Home	0.1556*	0.1023*	0.1610*	0.1587*	0.1447*
	Artificial Grass	-0.0275	-0.0683	0.0000	0.0000	0.0095
	Derby	0.0134	-0.0024	0.0000	0.0000	0.0333
	Weekday	-0.0222	0.0035	0.0000	0.0000	0.0113
	LogAttendance	0.0947*	0.1326*	0.0000	0.0000	0.0214
	Temperature	0.1156*	0.1284*	0.0000	0.0000	-0.008
	Precipitation	-0.07	-0.0952*	0.0000	0.0000	-0.0132
	Expected Points	0.4002*	0.3306*	0.4484*	0.4808*	0.4048*
	Difference in Rank	-0.3198*	-0.2272*	-0.5451*	-0.5339*	-0.4898*
Team	Red Card in Team	-0.0616	-0.0969*	-0.1113*	-0.1417*	-0.0817*
	Team AV Market Value	0.3666*	0.3672*	0.3097*	0.3322*	0.3147*
	Team CV Market Value	-0.0147	0.0121	0.0446	0.0457	0.0328
	Team HHI	0.0189	-0.018	-0.0184	0.0064	-0.0146
	Team AV Height	-0.2048*	-0.2155*	-0.1906*	-0.2059*	-0.1738*
	Team CV Height	-0.0280	-0.0337	-0.0066	-0.0289	0.0115
	Team AV Age	-0.0464	-0.0734	-0.1572*	-0.1533*	-0.1660*
	Team CV Age	0.0053	-0.0132	-0.0878*	-0.0761	-0.0715
	Team AV Match Experience	0.1423*	0.1073*	-0.007	0.0053	-0.0062
	Team CV Match Experience	-0.1226*	-0.1163*	-0.1341*	-0.1335*	-0.1067*

Note: *$p<0.05$; pairwise correlations for players with playing time at least 45 minutes; DV is dependent variable

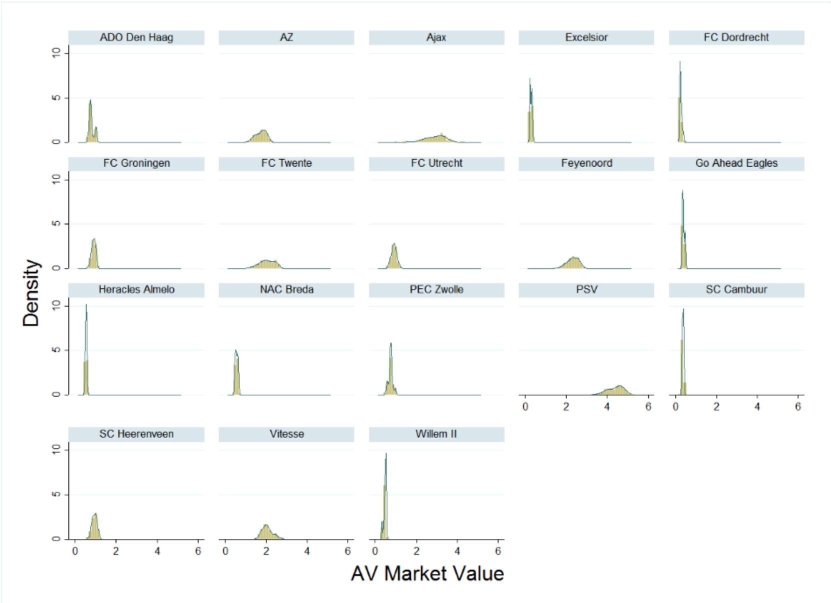


Figure E1: Density of AV Market Value by club

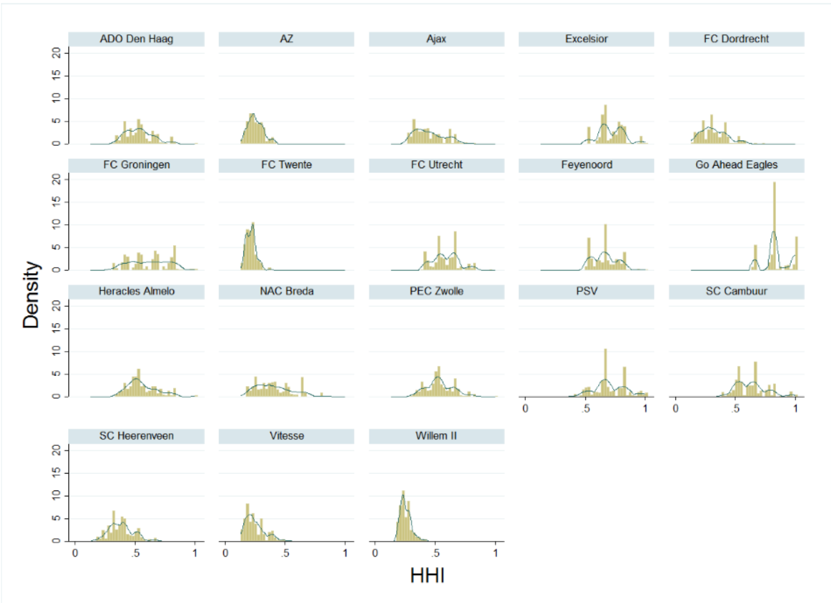


Figure E2: Density of HHI by club

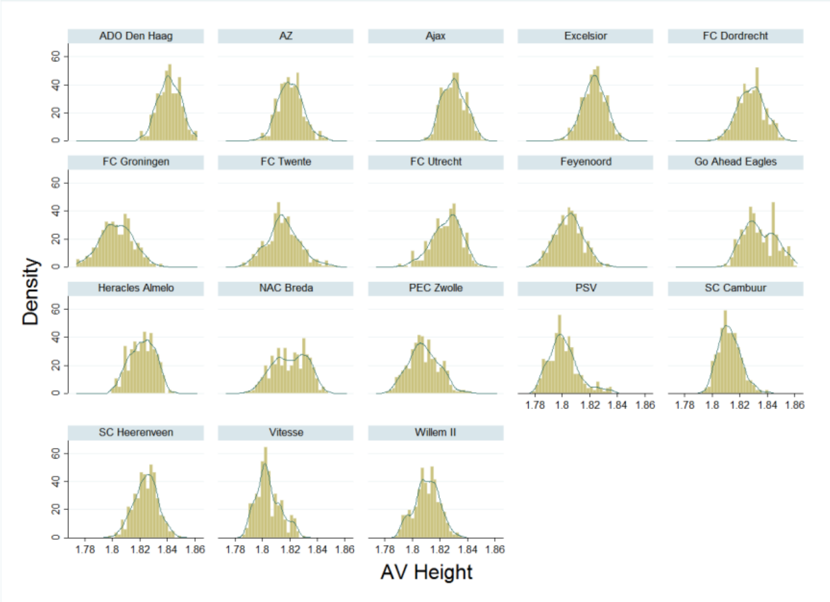


Figure E3: Density of AV Height by club

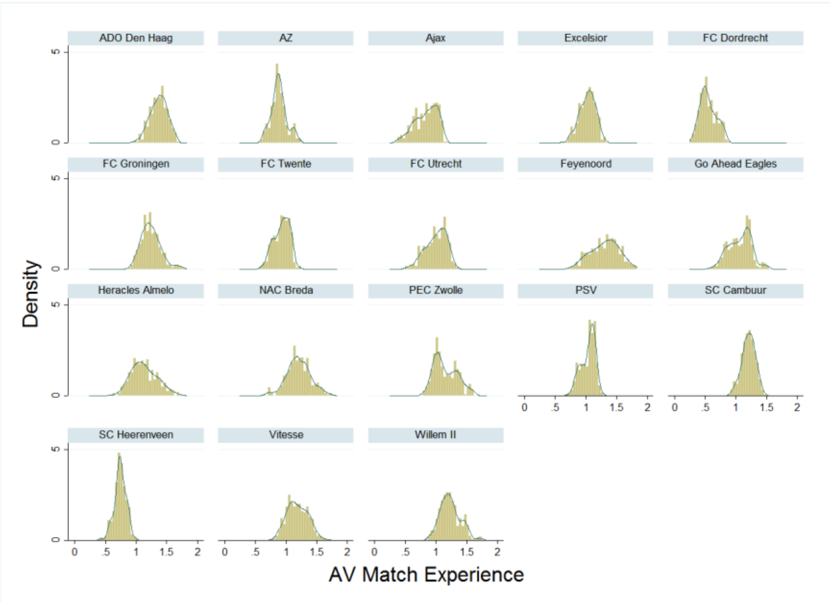


Figure E4: Density of AV Match Experience by club

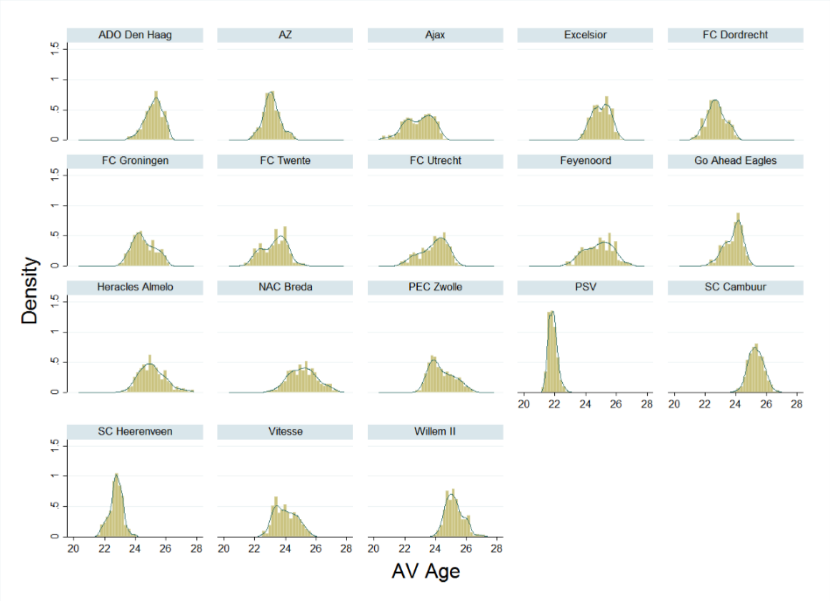


Figure E5: Density of AV Age by club

Table E4: Descriptive statistics by club for AV Market Value and CV Market Value

		AV Market Value				CV Market Value			
Club	N	Mean	SD	Min	Max	Mean	SD	Min	Max
ADO Den Haag	371	0.81	0.13	0.58	1.10	0.53	0.21	0.17	0.94
AZ	373	1.74	0.27	1.02	2.27	0.77	0.15	0.39	1.18
Ajax	374	2.91	0.54	1.03	4.30	0.70	0.18	0.33	1.47
Excelsior	374	0.28	0.05	0.17	0.40	0.64	0.11	0.26	0.96
FC Dordrecht	374	0.26	0.06	0.14	0.42	0.67	0.13	0.44	1.07
FC Groningen	374	0.90	0.11	0.60	1.12	0.63	0.10	0.44	0.92
FC Twente	375	2.02	0.38	1.10	2.71	0.75	0.16	0.40	1.17
FC Utrecht	373	0.91	0.14	0.54	1.28	0.82	0.11	0.57	1.07
Feyenoord	374	2.31	0.29	1.36	2.93	0.87	0.15	0.53	1.29
Go Ahead Eagles	373	0.39	0.06	0.30	0.54	0.51	0.09	0.28	0.70
Heracles Almelo	375	0.53	0.04	0.43	0.63	0.51	0.06	0.34	0.62
NAC Breda	373	0.55	0.06	0.41	0.69	0.56	0.09	0.30	0.74
PEC Zwolle	374	0.74	0.11	0.47	1.03	0.84	0.22	0.26	1.17
PSV	372	4.39	0.39	3.29	5.18	0.48	0.08	0.31	0.72
SC Cambuur	374	0.37	0.04	0.27	0.45	0.47	0.07	0.25	0.62
SC Heerenveen	372	0.94	0.12	0.61	1.21	0.79	0.12	0.47	1.03
Vitesse	374	2.05	0.26	1.48	2.78	0.62	0.10	0.37	0.82
Willem II	374	0.46	0.07	0.28	0.55	0.54	0.09	0.26	0.67
Total	6,723	1.25	1.10	0.14	5.18	0.65	0.18	0.17	1.47

Table E5: Descriptive statistics by club for HHI

		HHI			
Club	N	Mean	SD	Min	Max
ADO Den Haag	371	0.54	0.12	0.32	1.00
AZ	373	0.25	0.06	0.14	0.41
Ajax	374	0.45	0.12	0.27	0.82
Excelsior	374	0.72	0.11	0.52	1.00
FC Dordrecht	374	0.32	0.11	0.14	0.73
FC Groningen	374	0.62	0.16	0.32	1.00
FC Twente	375	0.22	0.04	0.14	0.38
FC Utrecht	373	0.59	0.11	0.40	0.97
Feyenoord	374	0.66	0.10	0.50	1.00
Go Ahead Eagles	373	0.84	0.10	0.66	1.00
Heracles Almelo	375	0.56	0.12	0.36	1.00
NAC Breda	373	0.41	0.15	0.19	0.82
PEC Zwolle	374	0.54	0.11	0.31	1.00
PSV	372	0.70	0.13	0.42	1.00
SC Cambuur	374	0.63	0.13	0.40	1.00
SC Heerenveen	372	0.38	0.10	0.18	0.67
Vitesse	374	0.25	0.08	0.13	0.52
Willem II	374	0.25	0.05	0.18	0.42
Total	6,723	0.50	0.21	0.13	1.00

Table E6: Descriptive statistics by club for AV Height and CV Height

		AV Height				CV Height			
Club	N	Mean	SD	Min	Max	Mean	SD	Min	Max
ADO Den Haag	371	1.841	0.008	1.821	1.862	0.031	0.004	0.019	0.038
AZ	373	1.821	0.009	1.794	1.848	0.034	0.002	0.027	0.039
Ajax	374	1.830	0.008	1.809	1.854	0.023	0.004	0.012	0.031
Excelsior	374	1.823	0.009	1.795	1.845	0.031	0.005	0.017	0.038
FC Dordrecht	374	1.829	0.010	1.799	1.861	0.036	0.004	0.026	0.045
FC Groningen	374	1.803	0.012	1.774	1.836	0.043	0.003	0.032	0.049
FC Twente	375	1.815	0.012	1.785	1.854	0.037	0.005	0.024	0.050
FC Utrecht	373	1.825	0.011	1.791	1.853	0.040	0.004	0.028	0.053
Feyenoord	374	1.805	0.010	1.777	1.836	0.037	0.004	0.026	0.052
Go Ahead Eagles	373	1.834	0.012	1.804	1.862	0.029	0.004	0.019	0.038
Heracles Almelo	375	1.822	0.009	1.801	1.844	0.042	0.004	0.028	0.048
NAC Breda	373	1.820	0.012	1.788	1.846	0.032	0.005	0.021	0.043
PEC Zwolle	374	1.808	0.011	1.783	1.843	0.037	0.004	0.026	0.046
PSV	372	1.801	0.011	1.780	1.836	0.036	0.004	0.024	0.044
SC Cambuur	374	1.813	0.008	1.793	1.840	0.028	0.003	0.018	0.035
SC Heerenveen	372	1.824	0.009	1.798	1.850	0.038	0.005	0.021	0.051
Vitesse	374	1.804	0.009	1.787	1.831	0.036	0.005	0.018	0.042
Willem II	374	1.811	0.009	1.789	1.835	0.032	0.004	0.019	0.040
Total	6,723	1.818	0.015	1.774	1.862	0.034	0.006	0.012	0.053

Table E7: Descriptive statistics by club for AV Match Experience and CV Match Experience

		AV Match Experience				CV Match Experience			
Club	N	Mean	SD	Min	Max	Mean	SD	Min	Max
ADO Den Haag	371	1.37	0.14	0.95	1.67	0.59	0.07	0.43	0.85
AZ	373	0.89	0.13	0.58	1.23	0.75	0.16	0.39	1.13
Ajax	374	0.84	0.18	0.37	1.17	0.90	0.15	0.58	1.48
Excelsior	374	1.03	0.13	0.61	1.31	0.64	0.10	0.44	0.90
FC Dordrecht	374	0.55	0.14	0.25	0.89	0.82	0.13	0.50	1.30
FC Groningen	374	1.24	0.16	0.87	1.79	0.62	0.14	0.36	0.99
FC Twente	375	0.93	0.13	0.57	1.23	0.63	0.12	0.41	0.97
FC Utrecht	373	1.00	0.17	0.51	1.35	0.90	0.15	0.56	1.36
Feyenoord	374	1.31	0.23	0.72	1.83	0.79	0.11	0.54	1.07
Go Ahead Eagles	373	1.07	0.19	0.57	1.52	0.99	0.21	0.47	1.45
Heracles Almelo	375	1.13	0.21	0.60	1.71	0.76	0.10	0.50	1.01
NAC Breda	373	1.21	0.19	0.63	1.72	1.00	0.16	0.53	1.36
PEC Zwolle	374	1.16	0.19	0.81	1.62	0.77	0.13	0.51	1.09
PSV	372	1.04	0.12	0.69	1.28	0.57	0.06	0.37	0.69
SC Cambuur	374	1.21	0.11	0.91	1.48	0.59	0.07	0.41	0.77
SC Heerenveen	372	0.74	0.10	0.41	1.00	0.79	0.08	0.58	1.04
Vitesse	374	1.19	0.17	0.79	1.66	0.65	0.11	0.37	0.87
Willem II	374	1.23	0.16	0.88	1.73	0.79	0.09	0.57	1.06
Total	6,723	1.06	0.26	0.25	1.83	0.75	0.18	0.36	1.48

Table E8: Descriptive statistics by club for AV Age and CV Age

		AV Age				CV Age			
Club	N	Mean	SD	Min	Max	Mean	SD	Min	Max
ADO Den Haag	371	25.21	0.57	23.51	26.33	0.11	0.02	0.06	0.15
AZ	373	23.11	0.55	21.75	24.60	0.13	0.02	0.08	0.18
Ajax	374	22.92	0.90	20.34	24.75	0.14	0.02	0.09	0.24
Excelsior	374	25.01	0.57	23.67	26.33	0.12	0.02	0.07	0.15
FC Dordrecht	374	22.73	0.60	21.26	24.11	0.10	0.02	0.07	0.15
FC Groningen	374	24.55	0.73	22.95	26.18	0.13	0.03	0.08	0.21
FC Twente	375	23.28	0.81	21.34	25.31	0.13	0.03	0.06	0.21
FC Utrecht	373	23.89	0.88	21.72	25.72	0.16	0.02	0.08	0.20
Feyenoord	374	24.86	0.91	22.72	26.94	0.17	0.02	0.14	0.23
Go Ahead Eagles	373	23.84	0.60	21.77	24.99	0.13	0.03	0.05	0.18
Heracles Almelo	375	25.11	0.85	23.01	27.83	0.16	0.01	0.09	0.19
NAC Breda	373	25.15	0.91	22.88	27.53	0.17	0.02	0.12	0.21
PEC Zwolle	374	24.40	0.81	22.81	26.58	0.15	0.03	0.10	0.21
PSV	372	21.86	0.29	21.28	22.95	0.09	0.02	0.06	0.16
SC Cambuur	374	25.32	0.50	23.80	26.78	0.12	0.01	0.09	0.15
SC Heerenveen	372	22.75	0.43	21.56	23.98	0.11	0.02	0.05	0.14
Vitesse	374	24.02	0.71	22.53	25.71	0.13	0.02	0.08	0.16
Willem II	374	25.23	0.57	23.92	27.13	0.13	0.02	0.07	0.19
Total	6,723	24.07	1.26	20.34	27.83	0.13	0.03	0.05	0.24

Appendix F: Baseline result with Passing%

Table F1: Parameter estimates baseline results individual Passing%

		(1)	(2)	(3)	(4)	(5)
Match	Home	0.010***	0.009**	0.009**	0.009**	0.004
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
	Artificial Grass	0.002	0.001	0.001	0.001	0.001
		(0.006)	(0.006)	(0.006)	(0.006)	(0.004)
	Derby	0.005	0.006	0.007	0.007	0.004
		(0.007)	(0.007)	(0.007)	(0.007)	(0.004)
	Weekday	0.002	0.002	0.003	0.003	0.003
		(0.009)	(0.010)	(0.010)	(0.010)	(0.006)
	LogAttendance	-0.003	-0.004	-0.004	-0.004	0.001
		(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
	Temperature	0.013***	0.013***	0.013***	0.013***	0.013***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.002)
	Precipitation	-0.012	-0.011	-0.012	-0.011	-0.014***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.004)
	Expected Points	0.003	0.001	0.001	0.001	0.013***
		(0.005)	(0.005)	(0.005)	(0.005)	(0.003)
	New Coach	-0.003	-0.002	-0.002		
		(0.008)	(0.009)	(0.009)		
Player	Playing time	0.019	0.019	0.019		
		(0.015)	(0.016)	(0.015)		
	Goalkeeper	0.052***	0.051***	0.052***	0.053***	0.053***
		(0.010)	(0.010)	(0.010)	(0.010)	(0.005)
	Defender	0.094***	0.094***	0.094***	0.095***	0.095***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.003)
	Midfielder	0.070***	0.070***	0.070***	0.071***	0.070***
		(0.007)	(0.007)	(0.007)	(0.006)	(0.004)
	Capped	0.012	0.012	0.012	0.011	0.006
		(0.010)	(0.010)	(0.010)	(0.010)	(0.004)
	Market Value	0.002	0.002	0.002	0.002	0.003**
		(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
	Age under 21	-0.011	-0.012	-0.012	-0.012	-0.014***
		(0.009)	(0.010)	(0.010)	(0.010)	(0.004)
	Age between 21 and 24	-0.015*	-0.015*	-0.015*	-0.014*	-0.018***
		(0.008)	(0.008)	(0.008)	(0.008)	(0.003)
	Age between 24 and 27	0.004	0.005	0.005	0.005	0.003
		(0.006)	(0.006)	(0.006)	(0.006)	(0.004)
	Dutch	-0.008	-0.010	-0.010	-0.010	-0.007**
		(0.008)	(0.007)	(0.008)	(0.008)	(0.003)
	Height	0.183***	0.171***	0.171***	0.173***	0.140***
		(0.050)	(0.050)	(0.050)	(0.050)	(0.022)
Team	Team Member Red Card		-0.012*	-0.013*	-0.013*	-0.012**
			(0.006)	(0.006)	(0.006)	(0.005)
	AV Market Value		0.008	0.006	0.006	0.009***
			(0.008)	(0.007)	(0.007)	(0.002)
	CV Market Value		0.009			
			(0.019)			
	HHI		-0.036**	-0.034**	-0.034**	0.009
			(0.015)	(0.016)	(0.015)	(0.006)
	AV Height		-0.183	-0.149	-0.148	-0.453***
			(0.213)	(0.205)	(0.203)	(0.090)
	CV Height		-0.446			
			(0.542)			
	AV Age		0.002	0.002	0.002	0.001
			(0.003)	(0.003)	(0.003)	(0.001)
	CV Age		-0.109***	-0.109**	-0.105**	0.047
			(0.037)	(0.049)	(0.049)	(0.044)
	Constant	0.402***	0.755	0.685	0.699	1.229***
		(0.093)	(0.449)	(0.431)	(0.430)	(0.186)
	Number of club	Yes	Yes	Yes	Yes	No

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1*. All estimates are based on 6,719 observations with playing time ≥ 45 minutes. Dependent variable is individual Passing%. Estimates in models (1), (2), (3) and (4) contain 18 club fixed effects with standard errors clustered by club, estimate in model (5) does not include any fixed effect.

Table F2: Parameter estimates for different variables of HHI

	(1) Player Level	(2) Team Level	(3) Player Level	(4) Team Level	(5) Player Level	(6) Team Level
HHI 20th	-0.008**	-0.002				
	(0.003)	(0.002)				
HHI 30th	-0.006	-0.004				
	(0.005)	(0.004)				
HHI 40th	-0.004	0.001				
	(0.004)	(0.004)				
HHI 50th	-0.006	-0.004				
	(0.005)	(0.005)				
HHI 60th	-0.012**	-0.007				
	(0.005)	(0.005)				
HHI 70th	-0.004	-0.006				
	(0.005)	(0.004)				
HHI 80th	-0.012**	-0.005				
	(0.005)	(0.005)				
HHI 90th	-0.007	-0.007				
	(0.005)	(0.005)				
HHI 100th	-0.015**	-0.007				
	(0.006)	(0.005)				
HHI defense			0.003	-0.001		
			(0.008)	(0.006)		
HHI midfield			0.000	-0.007*		
			(0.009)	(0.004)		
HHI attack			-0.003	-0.007		
			(0.009)	(0.007)		
HHI left					-0.013*	0.004
					(0.007)	(0.006)
HHI central					0.007	-0.001
					(0.005)	(0.005)
HHI right					-0.011	-0.015***
					(0.007)	(0.005)
Observations	6,719	612	6,719	612	6,719	612

Note: Robust standard errors in parentheses; ****p<0.01, **p<0.05, *p<0.1.* All estimates contain 18 club fixed effects. Models (1), (3) and (5) are based on a player level analysis with the individual Success Ratio as dependent variable and includes all the other variables as included in specification (4) of Table 5.1. Models (2), (4) and (6) are based on a team level analysis with the Team AV Success Ratio as dependent variable and includes all the other variables as included in specification (1) of Table 5.4. Standard errors are clustered by club. In Models (1) and (2), the number in the variable name represents a percentile. The corresponding upper values for the HHI dummies in the player level analysis are approximately 0.22 (HHI 10th; the reference category); 0.27; 0.34; 0.42; 0.50; 0.54; 0.64; 0.67; 0.81 and 1. The upper values for the HHI dummies in the team level analysis are approximately 0.21 (HHI 10th; the reference category); 0.26; 0.34; 0.43; 0.49; 0.55; 0.63; 0.69; 0.80 and 1. In models (3)-(6) the values for HHI are based on the subgroup of players, taking playing time into account, for the respective positions. The goalkeeper belongs to the defense.

Appendix G: Scatterplot with alternative number of points

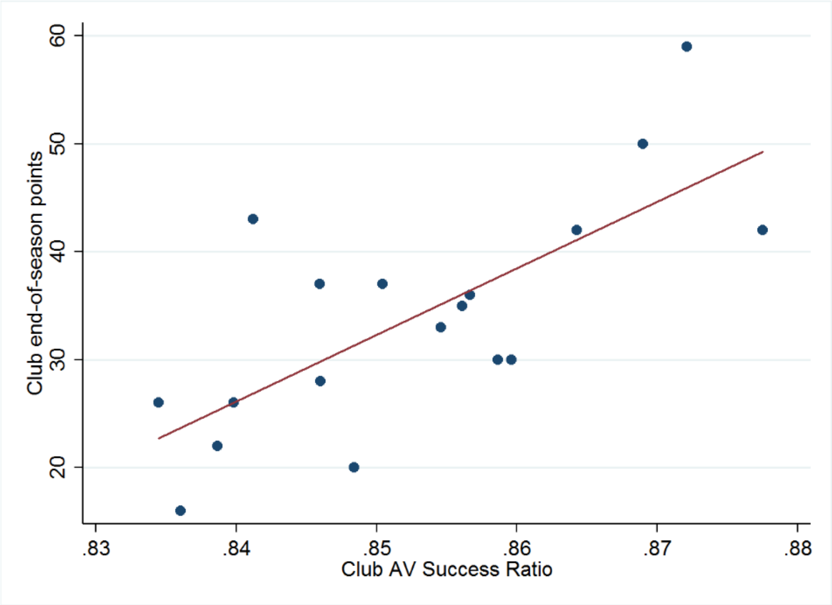


Figure G1: Scatterplot of club average Success Ratio and club end-of-season number of points (win 2 points, draw 1 point)

Appendix H: Results for simulations with alternative dependent variable

Table H1: Simulations of team performance

		Value	Team AV Passing%	Absolute difference compared to mean	% difference compared to mean
AV Market Value	Min	0.17	0.7773	-0.0149	-0.0188
	Mean-SD				
	Mean	1.25	0.7922	0	0
	Mean+SD	2.35	0.8073	0.0151	0.0191
	Max	4.85	0.8417	0.0495	0.0625
HHI	Min	0.14	0.8053	0.0131	0.0165
	Mean-SD	0.28	0.8000	0.0078	0.0099
	Mean	0.49	0.7922	0	0
	Mean+SD	0.70	0.7844	-0.0078	-0.0099
	Max	1.00	0.7732	-0.0190	-0.0240
AV Height	Min	1.79	0.7910	-0.0012	-0.0016
	Mean-SD	1.81	0.7918	-0.0004	-0.0005
	Mean	1.82	0.7922	0	0
	Mean+SD	1.83	0.7926	0.0004	0.0005
	Max	1.85	0.7934	0.0012	0.0016
CV Match Experience	Min	0.42	0.8034	0.0112	0.0141
	Mean-SD	0.59	0.7978	0.0056	0.0071
	Mean	0.76	0.7922	0	0
	Mean+SD	0.93	0.7866	-0.0056	-0.0071
	Max	1.34	0.7731	-0.0191	-0.0241
		Value	Points	Absolute difference compared to mean	% difference compared to mean
AV Market Value	Min	0.17	1.2090	0.0855	0.0761
	Mean-SD				
	Mean	1.25	1.1235	0	0
	Mean+SD	2.35	1.0364	-0.0871	-0.0775
	Max	4.85	0.8385	-0.2850	-0.2537
HHI	Min	0.14	0.9876	-0.1359	-0.1210
	Mean-SD	0.28	1.0420	-0.0815	-0.0726
	Mean	0.49	1.1235	0	0
	Mean+SD	0.70	1.2051	0.0815	0.0726
	Max	1.00	1.3216	0.1980	0.1763
AV Height	Min	1.79	1.4024	0.2789	0.2482
	Mean-SD	1.81	1.2165	0.0930	0.0827
	Mean	1.82	1.1235	0	0
	Mean+SD	1.83	1.0306	-0.0930	-0.0827
	Max	1.85	0.8447	-0.2789	-0.2482
CV Match Experience	Min	0.42	1.4164	0.2929	0.2607
	Mean-SD	0.59	1.2700	0.1464	0.1303
	Mean	0.76	1.1235	0	0
	Mean+SD	0.93	0.9771	-0.1464	-0.1303
	Max	1.34	0.6239	-0.4996	-0.4447
		Value	Goal Difference	Absolute difference compared to mean	% difference compared to mean
AV Market Value	Min	0.17	0.5284	0.2073	0.6457
	Mean-SD				
	Mean	1.25	0.3211	0	0
	Mean+SD	2.35	0.1099	-0.2112	-0.6576
	Max	4.85	-0.3700	-0.6911	-2.1522
HHI	Min	0.14	0.0651	-0.2560	-0.7971
	Mean-SD	0.28	0.1675	-0.1536	-0.4783
	Mean	0.49	0.3211	0	0
	Mean+SD	0.70	0.4747	0.1536	0.4783
	Max	1.00	0.6941	0.3730	1.1615
AV Height	Min	1.79	1.0679	0.7468	2.3257
	Mean-SD	1.81	0.5700	0.2489	0.7752
	Mean	1.82	0.3211	0	0
	Mean+SD	1.83	0.0722	-0.2489	-0.7752
	Max	1.85	-0.4257	-0.7468	-2.3257
CV Match Experience	Min	0.42	0.9421	0.6210	1.9340
	Mean-SD	0.59	0.6316	0.3105	0.9670
	Mean	0.76	0.3211	0	0
	Mean+SD	0.93	0.0106	-0.3105	-0.9670
	Max	1.34	-0.7383	-1.0594	-3.2992
		Value	Victory	Absolute difference compared to mean	% difference compared to mean
AV Market Value	Min	0.17	0.4826	0.0622	0.1479
	Mean-SD				
	Mean	1.25	0.4204	0	0
	Mean+SD	2.35	0.3571	-0.0633	-0.1507
	Max	4.85	0.2131	-0.2073	-0.4931
HHI	Min	0.14	0.3794	-0.0410	-0.0975
	Mean-SD	0.28	0.3958	-0.0246	-0.0585
	Mean	0.49	0.4204	0	0
	Mean+SD	0.70	0.4450	0.0246	0.0585
	Max	1.00	0.4802	0.0597	0.1421
AV Height	Min	1.79	0.5385	0.1181	0.2809
	Mean-SD	1.81	0.4598	0.0394	0.0936
	Mean	1.82	0.4204	0	0
	Mean+SD	1.83	0.3810	-0.0394	-0.0936
	Max	1.85	0.3023	-0.1181	-0.2809
CV Match Experience	Min	0.42	0.5409	0.1205	0.2866
	Mean-SD	0.59	0.4806	0.0602	0.1433
	Mean	0.76	0.4204	0	0
	Mean+SD	0.93	0.3602	-0.0602	-0.1433
	Max	1.34	0.2149	-0.2055	-0.4888

Note: Simulation based on results of models (2)-(4) in Table 5.4. In all these simulations we set the Home dummy at value one, while Artificial Grass, Derby and Weekday are set to zero. Also, Team Member Red Card is set at zero. All other variables are set at sample mean.

Bibliography

Abbott, A. and Collins, D. (2004) Eliminating the dichotomy between theory and practice in talent identification and development: considering the role of psychology. *Journal of Sports Sciences*, 22, 395-408.

Abbot, A., Collins, D., Martindale, R., and Sowerby, K. (2002) *Talent identification and development: An academic review.* A report for SportScotland by the University of Edinburgh. Edinburgh: sportscotland.

Addona, V., and Yates, P. A. (2010) A closer look at the relative age effect in the national hockey league. *Journal of Quantitative Analysis in Sports*, 6, 1-17.

Albuquerque, M. R., Fukuda, D. H., Da Costa, V. T., Lopes M. C., and Franchini, E. (2016) Do weight categories prevent athletes from the relative age effect? A meta-analysis of combat sports. *Sport Sciences for Health*, 12, 133-139.

Anderson, C., and Sally, D. (2013) *The numbers game; why everything you know about soccer is wrong*. New York: Viking/Penguin.

Angrist, J. D., and Krueger, A. B. (1991) Does compulsory school attendance affect schooling and earnings?. *Quarterly Journal of Economics*, 106, 979-1014.

Ashworth, J., and Heyndels, B. (2007) Selection bias and peer effects in team sports: the effect of age grouping on earnings of German soccer players. *Journal of Sports Economics*, 8, 355-377.

Audas, R., Dobson, S., and Goddard, J. (1999) Organizational performance and managerial turnover. *Managerial and Decision Economics*, 20, 305-318.

Augste, C. and Lames, M. (2011) The relative age effect and success in German elite U-17 soccer teams. *Journal of Sports Sciences*, 29, 983-987.

Bachan, R., Reilly, B., and Witt, R. (2008) The hazard of being an English football league manager: empirical estimates for three recent league seasons. *Journal of the Operational Research Society*, 59, 884-891.

Baimbridge, M. (1997) Match attendance at Euro 96: Was the crowd waving or drowning?. *Applied Economics Letters*, 4, 555-558.

Baimbridge, M., Cameron, S., and Dawson, P. (1996) Satellite television and the demand for football: A whole new ball game?. *Scottish Journal of Political Economy*, 43, 317-333.

Balduck, A., Buelens, M., and Philippaerts, R. (2010a) Short-term effects of midseason coach turnover on team performance in soccer. *Research Quarterly for Exercise and Sport*, 81, 379-383.

Balduck, A., Prinzie, A., and Buelens, M. (2010b) The effectiveness of coach turnover and the effect on home team advantage, team quality and team ranking. *Journal of Applied Statistics*, 37, 679-689.

Barnsley, R. H., and Thompson, A. H. (1988) Birthdate and success in minor hockey: the key to the NHL. *Canadian Journal of Behavioral Science*, 20, 167-176.

Barnsley, R. H., Thompson, A. H. and Barnsley, R. E. (1985) Hockey success and birth date: the relative age effect. *Canadian Association for Health, Physical Education, and Recreation*, 51, 23-28.

Barnsley, R. H., Thompson, A. H., and Stebelsky, G. (1991) "Born to play ball" the relative age effect and Major League Baseball. *Sociology of Sport Journal*, 8, 146-151.

Bartling, B., Brandes, L., and Schunk, D. (2015) Expectations as reference points: field evidence from professional soccer. *Management Science*, 61, 2646-2661.

Baxter-Jones, A. D. G., Helms, P., Maffulli, N., Bains-Preece, J. C., and Preece, M. (1995) Growth and development of male gymnasts, swimmers, soccer and tennis players: a longitudinal study. *Annals of Human Biology*, 22, 381-394.

Beck, N., Meyer, M. (2012) Modelling team performance: theoretical and empirical annotations on the analysis of football data. *Empirical Economics*, 43, 335-356.

Benz, M.-A., Brandes, L., and Franck, E. (2009) Do soccer associations really spend on a good thing? Empirical evidence on heterogeneity in the consumer response to match uncertainty of outcome. *Contemporary Economic Policy*, 27, 216-235.

Besters, L. M., Van Ours, J. C., and Van Tuijl, M. A. (2017) Outcome uncertainty, team quality and stadium attendance in Dutch professional football. *Mimeo.*

Bojke, C. (2007) The impact of post-season play-off systems on the attendance at regular season games, in: J. Albert and R. Koning (eds.), *Statistical Thinking in Sports*, Chapman and Hall/CRC, 179-202.

Borland, J., and Macdonald, R. (2003) Demand for Sport. *Oxford Review of Economic Policy*, 19, 478-502.

Börsch-Supan, A., and Weiss, M. (2016) Productivity and age: evidence from work teams at the assembly line. *Journal of the Economics of Ageing*, 7, 30-42.

Brandes, L., Franck, E., and Theiler, P. (2009) The effect from national diversity on team production: empirical evidence from the sports industry. *Schmalenbach Business Review*, 61, 255-246.

Bruinshoofd, A., and Ter Weel, B. (2003) Manager to go? Performance dips reconsidered with evidence from Dutch football. *European Journal of Operational Research*, 148, 233-246.

Bryson, A., Buraimo, B., and Simmons, R. (2011) Do salaries improve worker performance? *Labour Economics*, 18, 242-433.

Budzinski, O., and Pawlowski, T. (2017) The behavioral economics of competitive balance: theories, findings, and implications. *International Journal of Sport Finance*, 12, 109-122.

Buraimo, B. (2008) Stadium attendance and television audience demand in English league football. *Managerial and Decision Economics*, 29, 513-523.

Buraimo, B., Frick, B., Hickfang, M., and Simmons, R. (2015) The economics of long-term contracts in the footballers' labour market. *Scottish Journal of Political Economy*, 62, 8-24.

Buraimo, B., and Simmons, R. (2008) Do sports fans really value uncertainty of outcome? Evidence from the English Premier League. *International Journal of Sport Finance*, 3, 146-155.

Buraimo. B., and Simmons. R. (2015) Uncertainty of Outcome or Star Quality? Television Audience Demand for English Premier League Football. *International Journal of the Economics of Business*, 22, 449-469.

Cairns, J. A. (1987) Evaluating changes in league structure: The reorganization of the Scottish Football League. *Applied Economics*, 19, 259-275.

Cairns, J. A., Jennett, N., and Sloane, P. J. (1986) The economics of professional team sports: A survey of theory and evidence. *Journal of Economic Studies*, 13, 3-80.

Carter, N. (2006) *The Football Manager; a History*; London: Routledge.

Carter, N. (2007) 'Managing the media': The changing relationship between football managers and the media. *Sport in History*, 27, 217-240.

Chang, Y. Y., Dasgupta, S., and Hilary, G., (2010) CEO ability, pay and firm performance. *Management Science*, 56, 1633-1652.

Coates, D., and Humphreys, B. R. (2005) Novelty effects of new facilities on attendance at professional sporting events. *Contemporary Economic Policy*, 23, 436-455.

Coates, D., and Humphreys, B. R. (2012) Game attendance and outcome uncertainty in the National Hockey League. *Journal of Sports Economics*, 13, 364-377.

Coates, D., Humphreys, B. R., and Zhou, L. (2014) Reference-dependent preferences, loss aversion, and live game attendance. *Economic Inquiry*, 52, 959-973.

Cobley, S., Baker, J., Wattie, N., and McKenna, J. (2009) Annual age-grouping and athlete development: a meta-analytical review of relative age effects in sport. *Sports Medicine*, 39, 235-256.

Cox, A. (2015) Spectator demand, uncertainty of results, and public interest: evidence from the English Premier League. *Journal of Sports Economics*, 1-28.

Cumming, J., and Hall, C. (2002) Deliberate imagery practice: the development of imagery skills in competitive athletes. *Journal of Sports Sciences*, 20, 137-145.

Czarnitzki, D., and Stadtmann, G. (2002) Uncertainty of outcome versus reputation: Empirical evidence for the first German football division. *Empirical Economics*, 27, 101-112.

d'Addona, S., and Kind, A. (2014) Forced manager turnover in English soccer leagues: a long-term perspective. *Journal of Sports Economics*, 15, 150-79.

Delorme, N., Boiché, J., and Raspaud, M. (2010) Relative age effect in elite sports: methodological bias or real discrimination?. *European Journal of Sport Science*, 10, 91-96.

Delorme, N., and Raspaud, M. (2009) Is there an influence of relative age on participation in nonphysical sports activities? the example of shooting sports. *Journal of Sports Sciences*, 27, 1035-1042.

De Paola, M., and Scoppa, V. (2011) The effects of managerial turnover: evidence from coach dismissals in Italian soccer teams. *Journal of Sports Economics*, 13, 152-168.

Dhuey, E., and Lipscomb, S. (2008) What makes a leader? relative age and high school leadership. *Economics of Education Review*, 27, 173-183.

Dobson, S. M., and Goddard, J. A. (1992) The demand for standing and seated viewing accommodation in the English Football League. *Applied Economics*, 24, 1155-1163.

Dobson, S., and Goddard, J. (2011) *The Economics of Football*. 2nd edition. Cambridge: Cambridge University Press.

Doyle, J. R., Bottomley, P. A., and Angell, R. (2017) Tails of the traveling Gaussian model and the relative age effect: tales of age discrimination and wasted talent, *PLoS ONE*, 12.

Dudink, A. (1994) Birth date and sporting success. *Nature*, 368, 592.

Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993) The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363-406.

Falter, J. M., and Perignon, C. (2000) Demand for Football and intramatch winning probability: An essay on the glorious uncertainty of sports. *Applied Economics*, 32, 1757-1765.

Feddersen, A., Maenning, W., and Borcherding, M. (2006) The novelty effect of new soccer stadia: the case of Germany. *International Journal of Sport Finance*, 1, 174-188.

Forrest, D., Beaumont, J., Goddard, J., and Simmons, R. (2005) Home advantage and the debate about competitive balance in professional sports leagues. *Journal of Sports Sciences*, 23, 439-445.

Forrest, D., and Simmons, R. (2002) Outcome Uncertainty and Attendance Demand in Sport: The Case of English Soccer. *The Statistician*, 51, 229-241.

Forrest, D., and Simmons, R. (2006) New issues in attendance demand: The case of the English Football League. *Journal of Sports Economics*, 7, 247-266.

Forrest, D., Simmons, R., and Buraimo, B. (2005) Outcome uncertainty and the couch potato audience. *Scottish Journal of Political Economy*, 52, 641-661.

Franck, E., and Nüesch, S. (2010) The effect of talent disparity on team productivity in soccer. *Journal of Economic Psychology*, 31, 218-229.

Franck, E., Nüesch, S., and Pieper, J. (2011) Specific human capital as a source of superior team performance. *Schmalenbach Business Review*, 63, 376-392.

Furley, P., Memmert, D., and Weigelt, M. (2016) "How much is that player in the window? The one with the early birthday?" Relative age influences the value of the best soccer players, but not the best businesspeople, *Frontiers in Psychology*, 7, 84.

Gagné, F. (1993) Constructs and models pertaining to exceptional human abilities. In K. A. Heller, F. J. Mönks and A. H. Passow (eds.), *International handbook of research and development of giftedness and talent*, Oxford, Pergamon Press, 63-85.

Gagné, F. (2004) Transforming gifts into talents: the DMGT as a developmental theory. *High Ability Studies*, 15, 119-147.

Gagné, F. (2010) Motivation within the DMGT 2.0 framework. *High Ability Studies*, 21, 81-99.

Gangloff, K. A., Connelly, B. L., and Shook, C. L. (2016) Of scapegoats and signals: investor reactions to CEO succession in the aftermath of wrongdoing. *Journal of Management*, 42, 1614-1634.

García, J., and Rodríguez, P. (2002) The determinants of football match attendance revisited: Empirical evidence from the Spanish Football League. *Journal of Sports Economics*, 3, 18-38.

Gerhards, J., and Mutz, M. (2017) Who wins the championship? market value and team composition as predictors of success in the European football leagues. *European Societies*, 19, 223-242.

Gladwell, M. (2008) *Outliers; the story of success*. New York: Little, Brown.

Goddard, J., and Asimakopoulos, I. (2004) Forecasting Football Results and the Efficiency of Fixed-odds Betting. *Journal of Forecasting*, 23, 51-66.

Gould, E. D., and Winter, E. (2009) Interactions between workers and the technology of production: evidence from professional baseball. *The Review of Economics and Statistics*, 91, 188-200.

Grondin, S., Deshaies, P., and Nault, L. P. (1984) Trimestres de naissance et participation au hockey et au volleyball. *La Revue Que´becoise de l'Activite´ Physique,* 2, 97-103.

Haas, H., and Nüesch, S. (2012) Are multinational teams more successful? *The International Journal of Human Resource Management*, 23, 3105-3113.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2003) Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, 111, 465-497.

Hart, R. A., Hutton, J., and Sharot, T. (1975) A statistical analysis of association football attendances. *Applied Statistics*, 24, 17-27.

Helsen, W. F., Baker, J., Michiels, S., Schorer, J., Van Winckel, J., and Williams, M. (2012) The relative age effect in European professional soccer: did ten years of research make any difference?. *Journal of Sports Sciences*, 30, 1665-1671.

Helsen, W. F., Hodges, N. J., Van Winckel, J., and Starkes, J. L. (2000) The roles of talent, physical precocity and practice in the development of soccer expertise. *Journal of Sports Sciences*, 18, 727-736.

Helsen, W. F., Starkes, J. L., and Van Winckel, J. (1998) The influence of relative age on success and dropout in male soccer players. *American Journal of Human Biology*, 10, 791-798.

Helsen W. F., Starkes J. L., and Van Winckel J. (2000) Effect of a change in selection year on success in male soccer players. *American Journal of Human Biology*, 12, 729-35.

Helsen, W. F., Van Winckel, J., and Williams, A. M. (2005) The relative age effect in youth soccer across Europe. *Journal of Sports Sciences*, 23, 629-636.

Hentschel, S., Muehlheusser, G., and Sliwka, D. (2012) The impact of managerial change on performance. The role of team heterogeneity. *CESifo Working Paper 3950*.

Hoare D. (2000) Birthdate and basketball success: is there a relative age effect? proceedings of the pre-Olympic congress: sports medicine and physical education international congress on sports science; Sep, 7-13; Brisbane (QLD).

Howe, M. J. A., Davidson, J. W., and Sloboda, J. A. (1998) Innate talents: reality or myth?. *Behavioral and Brain Sciences*, 21, 399-442.

Huckman, R. S., Staats, B. R., and Upton, D. M. (2009) Team familiarity, role experience, and performance: evidence from Indian software services. *Management Science*, 55, 85-100.

Humphreys, B. R., and Zhou, L. (2015) The Louis-Schmelling paradox and the league standing effect reconsidered. *Journal of Sports Economics*, 16, 835-852.

Janssens, P. and Késenne, S. (1987) Belgian football attendances, *Tijdschrift voor Economie en Management*, 32, 305-15.

Jennett, N. (1984) Attendances, uncertainty of outcome and policy in Scottish League Football. *Scottish Journal of Political Economy*, 31, 176-198.

Jiménez, I. P., and Pain, M. T. G. (2008) Relative age effect in Spanish association football: its extent and implications for wasted potential. *Journal of Sports Sciences*, 26, 995-1003.

Kahane, L., Longley, N., and Simmons, R. (2013) The effects of coworker heterogeneity on firm-level output: assessing the impact of cultural and language diversity in the national hockey league. *The Review of Economics and Statistics*, 95, 302-314.

Kahn, L. (2000) The sports business as a labor market laboratory. *Journal of Economic Perspectives*, 14, 75-94.

Kahneman, D., and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, 47, 263-291.

Kandel, E., and Lazear, E. P. (1992) Peer pressure and partnerships. *Journal of Political Economy*, 100, 801-817.

Khanna, N., Poulsen, A. B. (1995) Managers of financially distressed firms: Villains or scapegoats?. *Journal of Finance*, 50, 919-940.

Koning, R. (2003) An econometric evaluation of the effect of firing a coach on team performance. *Applied Economics*, 35, 555-564.

Koning, R. H. (2007) Post-season play and league design in Dutch soccer, in: P. Rodriquez, S. Késenne and J. Garcia (eds.), *Governance and Competition in Professional Sports Leagues*, Oviedo: Ediciones de la Universidad de Oviedo, 191-215.

Koszegi, B., and Rabin, M. (2006) A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121, 1133-1165.

Krikendall, D. T. (2014) The relative age effect has no influence on match outcome in youth soccer. *Journal of Sport and Health Science*, 3, 273-278.

Kuper, S. and Szymanski, S. (2010) *Soccernomics*; London: Harper Collins.

Kuypers, T. (1997) The Beautiful Game? An Econometric Study of Audiences, Gambling and Efficiency in English Football, *Economics PhD Thesis, University College London*.

Lazear, E. P. (1999) Globalisation and the market for team-mates. *The Economic Journal*, 109, C15-C40.

Love, A., Kavazis, A. N., Morse, A., and Mayer Jr., K. C. (2013) Soccer-specific stadiums and attendance in Major League Soccer: investigating the novelty effect. *Journal of Applied Sport Management*, 5, 32-46.

Madalozzo, R., and Villar, R. B. (2009) Brazilian Football: What Brings Fans to the Game?. *Journal of Sports Economics*, 10, 639-650.

Maderer, D., Holtbrügge, D., and Schuster, T. (2014) Professional football squads as multicultural teams: cultural diversity, intercultural experience, and team performance. *International Journal of Cross Cultural Management*, 14, 215-238.

Mann, D. L., and Ginneken, P. J. M. A. (2017) Age-ordered shirt numbering reduces the selection bias associated with the relative age effect. *Journal of Sports Sciences*, 35, 784-790.

Martins, A. M., Cró, S. (2016) The demand for football in Portugal: new insights on outcome uncertainty. *Journal of Sports Economics*, 1-28.

McCarthy, N., and Collins, D. (2014) Initial identification & selection bias versus the eventual confirmation of talent: evidence for the benefits of a rocky road?. *Journal of Sports Sciences,* 32, 1604-1610.

McCarthy, N., and Collins, D., and Court, D. (2016) Start hard, finish better: further evidence for the reversal of the RAE advantage. *Journal of Sports Sciences*, 34, 1461-1465.

Miklós-Thal, J., and Ullrich, H. (2016) Career prospects and effort incentives: evidence from professional soccer. *Management Science*, 62, 1645-1667.

Muehlheusser, G., Scheemann, S., and Sliwka, D. (2016) The impact of managerial change on performance: the role of team heterogeneity. *Economic Inquiry*, 54, 1128-1149.

Musch, J., and Grondin, S. (2001) Unequal competition as an impediment to personal development: a review of the relative age effect in sport. *Developmental Review*, 21, 147-167.

Musch, J., and Hay, R. (1999) The relative age effect in soccer: cross-cultural evidence for a systematic discrimination against children born late in the competition year. *Sociology of Sport Journal*, 16, 54-64.

Neale, W. C. (1964) The peculiar economics of professional sports: A contribution to the theory of the firm in sporting competition and in market competition. *The Quarterly Journal of Economics*, 78, 1-14.

Nolan, J. E., and Howell, G. (2010) Hockey success and birth date: the relative age effect revisited. *International Review for the Sociology of Sport*, 45, 507-512.

Owen, D. (2013) Measurement of Competitive Balance and Uncertainty of Outcome, *University of Otago Economics Discussion Papers No. 1311*.

Papps, K. L., Bryson, A., and Gomez, R. (2011) Heterogeneous worker ability and team-based production: evidence from major league baseball 1920-2009. *Labour Economics*, 18, 310-319.

Pawlowski, T. (2013) Testing the Uncertainty of Outcome Hypothesis in European Professional Football: A stated Preference Approach. *Journal of Sports Economics*, 7, 247-266.

Pawlowski, T., and Anders, C. (2012) Stadium attendance in German professional football - The (un)importance of uncertainty of outcome reconsidered. *Applied Economics Letters*, 19, 1553-1556.

Pawlowski, T., Nalbantis, G., and Coates, D. (2017) Perceived game uncertainty, suspense and the demand for sport. *Economic Inquiry*, 1-20.

Peel, D. A., and Thomas, D. A. (1988) Outcome uncertainty and the demand for football: An analysis of match attendances in the English Football League. *Scottish Journal of Political Economy*, 35, 242-249.

Peel, D. A., and Thomas, D. A. (1992) The demand for football: Some evidence on outcome uncertainty. *Empirical Economics*, 17, 323-331.

Peel, D. A., and Thomas, D. A. (1996) Attendance demand: An investigation of repeat fixtures. *Applied Economics Letters*, 3, 391-394.

Pieper, J., Nüesch, S., and Franck, E. (2014) How performance expectations affect managerial replacement decisions. *Schmalenbach Business Review*, 66, 5-23.

Plug, E. J. S. (2001) Season of birth, schooling and earnings. *Journal of Economic Psychology*, 22, 641-660.

Poulsen, R. (2000) Should he stay or should he go? Estimating the effect of firing the manager in soccer. *Chance*, 13, 29-32.

Prat, A. (2002) Should a team be homogeneous? *European Economic Review*, 46, 1187-1207.

Rees, T., Hardy, L., Güllich, A., Abernethy, B., Côte, J., Woodman, T., Montgomery, H., Laing, S., and Warr, C. (2016) The Great British medalist project: a review of current knowledge on the development of the world's best sporting talent. *Sports Medicine*, 46, 1041-1058.

Rottenberg, S. (1956) The Baseball Players' Labor Market. *Journal of Political Economy*, 64, 242-258.

Salomo, S., and Teichmann, K. (2000) The relationship of performance and managerial succession in the German Premier Football League. *European Journal of Sport Management*, 7, 99-119.

Schreyer, D., Schmidt, S. L., and Torgler, B. (2016) Against all odds? Exploring the role of game outcome uncertainty in season ticket holders' stadium attendance demand. *Journal of Economic Psychology*, 56, 192-217.

Serrano, R., Garcia-Bernal, J., Fernandez-Olmos, M., and Espitia-Escuer, M. A. (2015) Expected quality in European football attendance: market value and uncertainty reconsidered. *Applied Economics Letters*, 22, 1051-1054.

Sherar, L. B., Baxter-Jones, A. D. G., Faulkner, R. A., and Russell K. W. (2007) Do physical maturity and birth date predict talent in male youth ice hockey players?. *Journal of Sports Sciences*, 25, 879-886.

Simmons, C., and Paull, G. C. (2001) Season-of-birth bias in association football. *Journal of Sports Sciences*, 19, 677-686.

Simon, H. A., and Chase, W. G. (1973) Skill in chess. *American Scientist*, 61, 394-403.

Simonton, D. K. (1999) Talent and its development: an emergenic and epigenetic model. *Psychological Review*, 106, 435-457.

Szymanski, S. (2003) The assessment: the economics of sport. *Oxford Review of Economic Policy*, 19, 467-477.

Tena, J. D., and Forrest, D. (2007) Within-season dismissal of football coaches: statistical analysis of causes and consequences. *European Journal of Operational Research*, 181, 362-373.

Ter Weel, B. (2011) Does manager turnover improve firm performance? Evidence from Dutch soccer. *De Economist*, 159, 279-303.

Thompson, A. H., Barnsley, R. H., and Battle, J. (2004) The relative age effect and the development of self-esteem. *Educational Research*, 46, 313-320.

Thompson, A. H., Barnsley, R. H., and Dyck, R. J. (1999) A new factor in youth suicide: the relative age effect. *Canadian Journal of Psychiatry*, 44, 82-85.

Toering, T. T., Elferink-Gemser, M. T., Jordet, G., Pepping, G-J., and Visscher C. (2011) Self-regulation of learning and relative age in elite youth soccer: International versus national level players. In *Self-regulation of learning and the performance level of youth soccer players (chapter 5 of the dissertation of T. T. Toering)*. Groningen: University of Groningen.

Toering, T. T., Elferink-Gemser, M. T., Jordet, G., and Visscher C. (2009) Self-regulation and performance level of elite and non-elite youth soccer players. *Journal of Sports Sciences*, 27, 1509-1517.

Torgler, B., and Schmidt, L. (2007) What shapes player performance in soccer? empirical findings from a panel analysis. *Applied Economics*, 39, 2355-2369.

Vaeyens, R., Lenoir, M., Williams, A. M., and Philippaerts, R. M. (2008) Talent identification and development programmes in sport, current models and future directions. *Sports Medicine*, 38, 703-714.

Vaeyens, R., Philippaerts, R. M., and Malina, R. M. (2005) The relative age effect in soccer: a match-related perspective. *Journal of Sports Sciences*, 23, 747-756.

Van Ours, J. C., and Stoeldraijer, L. (2011) Age, wage and productivity in Dutch manufacturing. *De Economist*, 159, 113-137.

Van Ours, J. C., and Van Tuijl, M. A. (2016) In-season head-coach dismissals and the performance of professional football teams. *Economic Inquiry*, 54, 591-604.

Van Rossum, J. H. A. (2006) Relative age effect revisited: findings from the dance domain. *Perceptual and Motor Skills*, 102, 302-308.

Verhulst, J. (1992) Seasonal birth distribution of West European soccer players: a possible explanation. *Medical Hypotheses*, 38, 346-348.

Vincent, J., and Glamser, F. D. (2006) Gender differences in the relative age effect among US Olympic development program youth soccer players. *Journal of Sports Sciences*, 24, 405-413.

Ward, A., Amason, A. C., Lee, P. M., and Graffin, S. D. (2011): “The scapegoat premium: a rational view of new CEO compensation”, in: M. A. Carpenter (ed.), *The handbook of research on top management teams*, Cheltenham, Edward Elgar Publishing Limited, 349-374.

Ward, P., Hodges, N. J., Starkes, J. L., and Williams, M. A. (2007) The road to excellence: deliberate practice and the development of expertise. *High Ability Studies*, 18, 119-153.

Wattie, N., Schorer, J., and Baker, J. (2015) The relative age effect in sport: a developmental systems model. *Sports Medicine*, 45, 83-94.

Williams, A. M., and Reilly, T. (2000) Talent identification and development in soccer. *Journal of Sports Sciences*, 18, 657-667.

Wilson, P., and Sim, B. (1995) The demand for semi-pro league football in Malaysia 1989–91: A panel data approach. *Applied Economics*, 27, 131-138.

Winter, E. (2004) Incentives and discrimination. *The American Economic Review*, 94, 764-773.